

Министерство общего и профессионального образования  
Российской Федерации

Уральский государственный университет им. А.М.Горького

**Л.Н.Мазур**

# **ИНФОРМАЦИОННЫЕ СИСТЕМЫ Теоретические проблемы**

**Екатеринбург 1997**

ББК Ч23я 73-1  
М139

Печатается по постановлению  
редакционно-издательского совета  
Уральского государственного  
университета им. А.М.Горького

**Мазур Л.Н. Информационные системы: Теоретические проблемы:**  
Учеб. пособие. Екатеринбург: УрГУ, 1997. 70 с.

Данное учебное пособие содержит основные сведения по теоретическим проблемам информационного поиска и информационных систем. Оно ориентировано на студентов, специализирующихся в области истории, архивоведения и документоведения.

Рецензенты: доценты П.И.Килин,  
Н.А.Гайдамакин

ISBN 5-230-06786-1

© Л.Н.Мазур, 1997

## Оглавление

<b>ВВЕДЕНИЕ .....</b>	<b>4</b>
<b>ГЛАВА 1. ИНФОРМАЦИЯ И ЕЕ РОЛЬ В ОБЩЕСТВЕ .....</b>	<b>6</b>
<b>ГЛАВА 2. НАУЧНО-ИНФОРМАЦИОННАЯ ДЕЯТЕЛЬНОСТЬ .....</b>	<b>17</b>
<b>ГЛАВА 3. ИНФОРМАЦИОННЫЕ СИСТЕМЫ .....</b>	<b>31</b>
3.1. Виды информационных систем.....	31
3.2. Информационный массив.....	37
3.3. Технические носители информации.....	39
3.4. Техническое оснащение информационных систем.....	41
3.5. Информационно-поисковые языки.....	43
3.6. Показатели эффективности информационной системы.....	51
<b>ГЛАВА 4. ИНТЕГРАЛЬНЫЕ И ЭКСПЕРТНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ .....</b>	<b>52</b>
<b>ГЛАВА 5. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В АРХИВНОМ ДЕЛЕ.....</b>	<b>61</b>
<b>СПИСОК ЛИТЕРАТУРЫ .....</b>	<b>69</b>
<b>СПИСОК СОКРАЩЕНИЙ .....</b>	<b>70</b>

## ВВЕДЕНИЕ

В середине текущего столетия появились публикации, в которых в числе причин, могущих привести цивилизацию к гибели, называлось перепроизводство информации. Это явление получило название информационного кризиса (объемы информации, обрушившиеся на обыкновенного человека, возросли многократно, и рост их приобрел лавинообразный характер). Ежедневно в мире издается в среднем около 100 печатных листов текста в расчете на одного специалиста, работающего в узкой отрасли науки или техники. При таком обилии информации ученым все сложнее справиться с ней, нередко дублирование в исследованиях. По подсчетам ученых США от 10 до 20% выполненных научно-технических работ являются прямым дублированием проведенных ранее исследований, причем на это тратится от 30 до 80% рабочего времени ученых.

Поток информации, обрушивающийся на человека, приводит к двум последствиям: во-первых, ему все труднее ориентироваться в этом информационном море, найти нужные сведения для работы и удовлетворения иных информационных потребностей. Во-вторых, такая ситуация приводит к постоянному росту потерь огромных информационных богатств, т.е. к возникновению информационного голода.

Причиной информационного кризиса можно считать несоответствие применяемых в обществе технологий сбора, хранения и поиска информации многократно возросшим объемам информации, которые необходимы для дальнейшего развития общества. Поэтому очевидна актуальность задачи совершенствования средств и методов информационного обеспечения. Только широкое внедрение современной организационной и вычислительной техники, освоение современных способов поиска и переработки информации позволит разрешить ситуацию информационного кризиса.

Мы живем в эпоху, которую принято называть информационным обществом. Информатизация – это глобальный процесс, она охватывает все сферы жизни человека и приводит к тому, что на первый план выдвигается процесс производства и использования информации. Информация – вот основное средство производства. Например, чтобы увеличить объемы промышленного производства в 2 раза, производство информации необходимо увеличить в 4 раза, а для 10-кратного увеличения материального производства объем информации должен возрасти в 100 раз.

Основным признаком информационного общества является общедоступность информации и знаний, а для этого необходимо наличие общедоступных информационных систем, баз данных. Но мало их создать, они должны быть востребованы обществом, т.е. общество и каждый человек должны обладать информационной культурой.

Сегодня необходимо осознать, что любая деятельность в современном обществе – научная, педагогическая, управленческая, коммерческая – невозможна без освоения приемов и способов рациональной работы с информацией. Важнейшими признаками информационной культуры являются: а) умение ориентироваться в потоке информации; б) способность правильно сформулировать свои информационные запросы; в) навык составления вторичных документов; г) отношение к информационным службам. Общим элементом информационной культуры является понимание возможностей современных информационных систем и умение использовать эти возможности для своей практической деятельности. Вместе с тем социологические обследования, проведенные американскими учеными, показывают, что даже при высоком уровне организации научно-информационной деятельности только 60% потенциальных потребителей информации пользуются услугами информационных служб. Основной причиной этого является низкая информационная культура.

Внедрение автоматизированных информационных систем во все сферы деятельности человека, в том числе и такие, как архивное, библиотечное, музейное дело, образование, научные исследования, приобретает все более нарастающие темпы. Компьютеры и информационные технологии могут и должны облегчить так называемый черновой труд научных работников, который связан с поиском необходимой информации, их оперативной обработкой. Все это позволит ученым сосредоточиться на анализе процессов и явлений, сопоставлении наблюдений и фактов и получении новых научных результатов, приведет к существенному сокращению сроков выполнения научных работ. Таким образом, компьютерная техника и автоматизированные информационные системы — это один из важнейших компонентов нового информационного общества.

## **Глава 1. Информация и ее роль в обществе**

Информацию нередко сравнивают с воздухом, без которого стала бы невозможна жизнь общества. И действительно, научно-технический прогресс, развитие современного производства, усложнение общественных процессов и структур, бурная политическая жизнь – все это питается информационными потоками и в свою очередь порождает новую информацию, возникающую в порядке обычного отражения событий и явлений, а также как результат направленной научно-исследовательской деятельности. Информация выполняет в обществе коммуникативные, управленческие, научные, познавательные, учебно-воспитательные и другие функции, которые не требуют дальнейшей расшифровки.

Для того, чтобы любая система, в том числе социальная, нормально функционировала, необходим интенсивный и бесперебойный обмен информацией между ее элементами. Отсутствие организованных информационных потоков или нарушение их движения неизменно ведет к разрушению устойчивости системы, невозможности эффективно управлять ею, планировать и прогнозировать ее дальнейшее развитие. Правильно построенное информационное обеспечение жизнедеятельности общества – одна из важнейших основ организованности социальной системы, залог ее выживаемости.

Для успешного социально-экономического развития необходимо иметь технические и материальные средства, человеческие ресурсы и, конечно же, достаточную научно-техническую и социальную информацию. Бурный XX в. принес нам четкое осознание общественной ценности информации, которая превратилась в поистине стратегический ресурс человечества. При этом она в отличие от физических ресурсов обладает уникальным свойством, которое можно назвать неистощимостью – при потреблении информации ее объем не уменьшается, а увеличивается.

Вместе с тем информация может быть не только полезной, но и вредной, что связано с неизбежным возрастанием в информационных потоках доли бесполезной информации, а нередко и прямой дезинформации. Поэтому она может терять свою функцию внесения упорядоченности в общественные процессы, а информационный кризис становится прямой угрозой нормальному развитию человечества.

Как научное понятие информация стала изучаться не так давно. В 1948 г. появились две фундаментальные работы – книги Н.Винера и

К. Шеннона<sup>1</sup>, в которых известное уже давно слово «информация» стало наполняться новым научным содержанием. В дальнейшем понятие информации разрабатывалось на протяжении десятилетий в рамках разных наук. Было много попыток дать общее единое определение, но они до сих пор не увенчались успехом, и причиной этого служат многоплановость и сложность рассматриваемого понятия, его всеобъемлющий характер. Если обобщить опыт изучения информации почти за полвека, то можно сделать вывод, что это общеполитическая категория, стоящая в ряду с такими понятиями, как пространство, время, движение, энергия.

К настоящему времени написано огромное количество статей, книг, диссертаций, посвященных изучению сущности информации и ее роли в обществе, анализу соотношения информации и мышления. Наиболее обобщенные результаты философского осмысления понятия «информация» представлены в работах А.Д.Урсула<sup>2</sup>. Не вдаваясь в особенности дискуссий философских школ, отметим, что взгляды ученых на природу и сущность этого явления разделились, и их можно свести к двум точкам зрения. Одну из них выражает философская школа А.Д.Урсула. Он считает, что информация является всеобщим свойством материи – от простейших неорганических форм до человеческого общества – и представляет собой отраженное разнообразие. В 80-е гг. сформировалась и стала преобладающей точка зрения группы философов, которые утверждали, что информация возникает лишь в высокоорганизованной материи и связана с процессом управления.

Академический философский словарь предлагает следующее определение информации:

1. Сообщения, сведения о чем-либо, передаваемые людьми.
2. Уменьшаемая, снимаемая неопределенность в результате получения сообщения.
3. Сообщение, неразрывно связанное с управлением.
4. Отражение разнообразия в любых объектах и процессах неживой и живой природы.

Таким образом, понятие «информация» предполагает: 1) наличие источника и потребителя информации и 2) полезность сведений, сообщений.

---

<sup>1</sup> См.: Винер Н. Кибернетика, или Управление и связь в животном и машине. 2-е изд. М., 1968; Шеннон К. Работы по теории информации и кибернетике. М., 1963.

<sup>2</sup> См., напр.: Урсул А.Д. Природа информации: Философский очерк. М., 1968; Он же. Проблемы информации в современной науке. М., 1975; и др.

С позиций управления под информацией понимаются сведения о системе и внешней среде, используемые для постановки и реализации управленческих задач. С точки зрения информационных систем, вполне допустимо использовать более узкое понятие информации как сведений, являющихся объектом хранения, передачи и преобразования.

Информация классифицируется на следующие типы (см. схему).



Под социальной информацией принято понимать информацию, циркулирующую в человеческом обществе. Ей противостоит несоциальная информация как признание всеобщности этого понятия, присущей не только социуму, но и природе в целом, т.е. это сведения, не реализованные еще в знаниях человечества, но присущие материи. Социальная информация в свою очередь подразделяется на семантическую и несемантическую. Семантическая (смысловая) информация представлена в языковых символах, в то время как к несемантической относят главным образом информацию, воспринимаемую органами чувств на уровне ощущений. И в свою очередь семантическая информация разбивается на научную и ненаучную. Ненаучная информация вырабатывается и реализуется в бытовой сфере, художественном творчестве, религии и т.д. Научная информация имеет свои специфические источники возникновения, которые связаны в первую очередь с научно-исследовательской деятельностью.

Научная информация – это получаемая в процессе познания логическая информация, которая адекватно отражает явления и законы природы, общества и мышления и используется в общественно-



исторической практике<sup>3</sup>. Приведенное определение содержит важнейшие характеристики сущности научной информации, которые отличают ее от любой другой. Во-первых, подчеркнем тот момент, что источником научной информации является познание как определенный вид деятельности человека и общества в целом, реализуемый не только в научно-исследовательской работе, но и общественной практике, т.е. в производственной, хозяйственной, общественной деятельности. Познание нацелено в первую очередь на проникновение в сущность окружающих объектов, явлений, получение полноценного знания о их природе и использование его в своих целях.

Во-вторых, научная информация имеет логическую форму выражения, т.е. научной она становится только тогда, когда подвергнута обработке и обобщению абстрактно-логическим мышлением. Как правило, средством ее выражения является язык науки и ее понятийный аппарат, который реализуется через систему суждений, умозаключений, теорий, гипотез.

Третий необходимый признак научной информации – адекватное отражение ею явлений и законов природы, общества и мышления. Под адекватностью отражения здесь понимаются истинные знания о каком-либо явлении или объекте реального мира, которые определяются уровнем развития науки. В процессе познания люди могут получать логическую информацию, которая дает искаженное представление об окружающем мире, в том числе это могут быть ошибочные теории и гипотезы. Но ошибочность таких гипотез и теорий не свидетельствует об их ненаучности. Гипотеза или теория является научной, если не запрещает систематическое изучение, обобщение и проверку на практике ее основных положений. В конечном итоге доказательство ошибочности есть также научный результат.

Наконец, четвертый признак научной информации заключается в том, что она должна быть использована в общественно-исторической практике. Это обеспечивает проверку истинности научных сведений и препятствует отнесению к научной информации общеизвестных или устаревших истин, научной фантастики и т.д.

Исходя из рассмотренного определения научной информации, охарактеризуем ее виды. Классификация научной информации может строиться исходя из различных оснований, среди которых наиболее значимыми являются следующие (см.табл.):

- по назначению;

---

<sup>3</sup> См.: Михайлов А.И., Черный А.И., Гиляревский Р.С. Научные коммуникации и информатика. М., 1976. С.73.

- по типу передаваемой информации;
- по способу распространения;
- по степени аналитико-синтетической обработки;
- по области получения (использования).

Этим многообразие видов научной информации не ограничивается, так как мы можем дополнить вышеуказанный перечень и другими основаниями классификации, например, по форме представления в научной информации выделяют визуальную; звуковую; тактильную (осязательную). В свою очередь визуальная информация может подразделяться на текстовую и графическую. Таким образом системы классификации научной информации по различным основаниям дают нам видимое представление о многообразии форм ее проявления и использования.

Для того, чтобы иметь более систематизированное представление о классификации научной информации посмотрим таблицу.

*Таблица*

**Виды научной информации**

Основание классификации	Наименование признака	Примечание
1. Назначение	Массовая Специальная	Предназначенная для всех, а не только для специалистов Предназначенная только для специалистов
2. Тип передаваемой информации	Документальная Фактографическая	Информация о научных документах, их местонахождении, содержании Идеи, факты, данные, извлеченные из научных документов
3. Способ распространения информации	Опубликованная Неопубликованная	Представлена в тиражном издании Не рассчитана на широкое распространение (научные отчеты, диссертации, депонированные рукописи и т.д.)
4. Степень аналитико-синтетической обработки информации	Первичная Вторичная	Непосредственные результаты научно-исследовательской и опытно-конструкторской работы Результат аналитико-синтетической переработки первичной информации; сведения о документах, содержащих первичную информацию
5. Область получения (использования)	Астрономическая, биологическая, идеологическая, историческая, медицинская, сельскохозяйственная, техническая, управленческая и т.д.	

Научная информация обладает определенной структурой, в которой можно выделить формальный и содержательный аспекты. По содержанию в структуре научной информации можно выделить научные данные и научные знания. К научным данным относятся сведения, полученные в процессе чувственного познания и еще не подвергнутые обработке абстрактно-логическим мышлением. Это своеобразное сырье научной информации. Под научными знаниями понимают информацию, представленную уже в наиболее обобщенном систематизированном виде, выраженную через понятия, суждения и умозаключения. Следовательно, научные знания представляют лишь часть научной информации, к которой относятся формулировка и проверка гипотез, концепций, законов, построение теорий – это наиболее высокий уровень обобщения, а также систематизация, обобщение, проверка на достоверность научных фактов.

Одной из форм организации данных является база данных. Под базой данных принято в настоящее время понимать организованную специальным образом совокупность данных, отображающую свойства и состояние конкретных объектов. Выделяются два аспекта баз данных: инфологический и даталогический.

Инфологический аспект включает вопросы смыслового содержания информации вне зависимости от способов ее представления и характеризует предметную область базы данных, т.е. ту часть реального мира, информационные потребности которого реализуются конкретной базой данных. Даталогический аспект включает вопросы представления и организации, правила смысловой интерпретации и преобразования данных.

Некоторые авторы, например В.И. Сифоров, предлагают более подробную структуру научной информации, выделяя следующие структурные элементы: 1) информация о научных фактах; 2) информация о научных гипотезах, теориях, концепциях частного характера; 3) информация, охватывающая систему фактов, гипотез, концепций, теорий и законов, которые составляют основу данной науки или области знания и 4) информация мировоззренческого уровня, отражающая общие подходы к познанию мира<sup>4</sup>.

Нужно отметить, что содержательная структура научной информации достаточно условна, так как одни и те же фрагменты

---

<sup>4</sup> См.: Сифоров В.И. Научная информация и повышение эффективности научных исследований // Управление, планирование и организация научных и технических исследований. М., 1970. Т.3. С.291-292.

информации могут относиться к разным уровням и очень трудно бывает провести четкую линию между фактом и знанием.

Формальная структура также иерархична. Первичная единица информации – сигнал. В технических системах используют сигналы самой различной природы (электрические, пневматические). В сигнальных системах типа человек—человек сигналами служат звук, буква, слово, фраза. Промежуточным структурным элементом информации служит знак (буквенный, цифровой или любой другой). Система знаков создает образ обозначаемого предмета.

Таким образом, мы можем представить схему трансформации информации при ее движении от объекта к субъекту: предмет – образ – знак – сигнал.

Со знаком, как со структурной единицей и формой выражения информации, тесно связано понятие кода. Код – это образ знака или слова естественного языка, выраженный в символах какого-либо искусственного языка. Вся информационная техника развивается в направлении поисков и создания искусственных языков – знаковых и кодовых систем, обеспечивающих наиболее эффективное представление, хранение и передачу информации.

Важная структурная единица информации – языки. Выделяют два больших класса языков: естественные и искусственные. Искусственные информационные языки делятся на следующие виды: классификационные, дескрипторные, алгоритмические, машинные.

Более крупная единица информации – информационная модель – представляет собой систему знаков, символов, слов, предложений искусственного или естественного языка, имитирующих объект в целом или его отдельные части. Известны 4 основные группы моделей: абстрактные (знаковые или графические), наглядные (изображения), предметные (макеты) и комбинированные.

Содержательный и формальный аспекты структуры научной информации также находятся в определенной взаимосвязи друг с другом. Каждый из названных выше содержательных классов научной информации тяготеет к определенным видам научных документов и отражается в формальной структуре.

Важной проблемой является изучение свойств информации, которые до сих пор не получили полного раскрытия в работах философов и специалистов. Из всех свойств, присущих научной информации, мы рассмотрим только некоторые, наиболее важные для построения информационных систем.

1. Неотрывность научной информации от физического носителя. Сущность этого свойства отражает тот факт, что информация не может существовать вне той или иной материальной оболочки. Она может быть закреплена на бумаге, магнитной ленте, папирусе либо в мозгу человека.

2. Языковая природа научной информации. Научная информация есть информация семантическая (понятийная) и использует языковую форму выражения. При этом под языком мы понимаем знаковую систему любой природы, которая служит средством общения между людьми, мышления и выражения. Язык может быть естественным и искусственным. Научная информация в чистом виде не существует в отрыве от языка, так как мышление тесно связано с языком, и только благодаря языковой природе научной информации возможно ее фиксирование на разных физических носителях для передачи в пространстве и времени. Языковую природу имеет не только научная информация, но и все другие виды социальной информации, за исключением некоторых видов эстетической информации.

3. Независимость научной информации от языка и физического носителя. Содержательный аспект научной информации не зависит от того, на каком языке она выражена и на каком физическом носителе закреплена. Смысл теории вероятности не изменится, если она будет изложена на русском, английском языках или в виде математических формул, написана будет на бумаге или магнитном носителе – любая из этих записей будет содержать семантически равноценную научную информацию. Однако это не универсальное свойство любой информации. Так, например, восприятие художественной информации во многом зависит от языка оригинала, формы ее выражения. Человек получает неодинаковую информацию, если читает трагедию «Гамлет» на английском языке или в переводе, смотрит разные постановки ее в театре, кино или слушает по радио.

4. Кумулятивность научной информации. Это свойство тесно связано с преемственностью науки и ее интернациональным характером. Достижения ученых всего мира и всех предшествующих поколений являются тем фундаментом, на котором продолжают свои исследования современные ученые. Естественно, что каждое поколение ученых занято не только получением новых знаний, но и целенаправленной работой по систематизации, оценке и обобщению научной информации, созданной ранее. Это возможно благодаря свойству кумулятивности, т.е. способности более краткого и обобщенного изложения первоначально созданного объема информации. Со временем все частное,

незначительное отсеивается, а главное получает более простое и краткое изложение. В точных и естественных дисциплинах, где язык науки более формализован, кумулятивность выражена сильнее. Нередко в одной формуле здесь заключен смысл, для изложения которого могут понадобиться тома книг. В общественных и гуманитарных науках это свойство проявляется слабее. Со свойством кумулятивности тесно связаны явления концентрации научной информации и перехода ее ко все более высоким уровням абстракции.

5. Независимость научной информации от ее создателей. Научная информация в отличие от художественной сразу же после создания приобретает известную самостоятельность и свободу от ее создателя и от произведений, в которых она впервые появилась. Причем, с течением времени степень этой свободы становится все больше. Это связано с тем, что форма выражения научных истин не играет особой роли в их дальнейшем использовании. Науке присущ общественный характер, знание есть достояние всех желающих к нему прикоснуться. Поэтому права авторства относятся к одной из болезненных проблем, также как и утверждение приоритета в открытиях.

6. Полезность научной информации. Данное свойство крайне субъективно, так как критерием полезности той или иной информации выступают потребности человека, информационной системы или общества в целом. Учет доли полезной информации является одним из важнейших показателей эффективности деятельности любой информационной системы.

7. Полнота научной информации. Мечтой любого ученого является владение полной информацией об изучаемом объекте. Однако мечта эта недостижима, потому что познание бесконечно. Поэтому мы можем говорить только о степени полноты информации, которая зависит от источников получения данных, методов исследования и многих других факторов. Специфика исторической информации состоит в том, что ее полнота уменьшается по мере удаленности от нас исторической эпохи. Вместе с тем степень полноты должна быть достаточной, чтобы достигнуть достоверных результатов исследования.

8. Достоверность. Итак, достоверность и полнота – два тесно взаимосвязанных свойства. Под достоверностью мы понимаем адекватное отражение объективного мира, законов его развития в фактах, представлениях и суждениях исследователей. Чем меньше полнота, тем меньше достоверность информации. Однако связь между ними не прямая, а логическая. Первые разрозненные сведения мало о чем говорят, по достижении определенного уровня полноты, когда

количество данных становится достаточным для проведения анализа и формулирования обоснованных выводов, наступает момент, когда количество исходной информации напрямую влияет на ее достоверность. При достижении следующего критического уровня новые факты становятся избыточными и не добавляют ничего принципиально нового. Наличие таких критических уровней полноты – минимального и максимального – делает принципиально возможным существование науки и познавательной деятельности как таковых. Нередко исследователям приходится иметь дело с преднамеренным искажением информации (дезинформацией). Особенно актуальна подобная проблема для общественных наук. Поэтому важное значение приобретает разработка методов доказательства достоверности информации, которые реализуются через такие формы, как логический анализ, контроль, эксперимент и т. д.

9. Старение научной информации. Мы часто говорим, что те или иные сведения устарели, подразумевая их неверность, неистинность, которые выявились по прошествии какого-либо времени в ходе научной и общественной практики. История науки имеет много примеров, когда устаревшие теории и концепции тормозили ее развитие, и наоборот, открытия Архимеда, Ньютона или Ломоносова трудно назвать устаревшими. Полностью устаревает лишь та информация, которая оказывается неверной, перестает адекватно отражать явления и закономерности реального мира. Изучение свойств информации позволило сделать следующие выводы: наибольшему старению подвержена информация, заключенная в гипотезах, частных концепциях и теориях, а наименьшему – в фактах, первичных данных, а также в законах, образующих основы определенной науки, т.е. научные факты и основные положения научных дисциплин опровергаются гораздо реже. Из источников информации с наибольшим недоверием следует относиться к периодической печати, в частности к газетной информации, которая стареет очень быстро.

10. Рассеяние научной информации. Это одна из важнейших проблем науки, которая стала краеугольным камнем научно-информационной деятельности, главной целью ее является как раз смягчение последствий рассеяния научной информации. Свойство рассеяния тесно связано с дискретностью информации, кумулятивностью, независимостью от создателя и старением. Научная информация отражена в различных публикациях: журналах, монографиях, статьях. Каждый автор по своему формулирует изучаемую проблему, уточняет объект исследования. Наиболее зримо со

свойством рассеяния мы сталкиваемся, когда встает необходимость выявить степень изученности той или иной проблемы. Даже для написания небольшого реферата исследователь привлекает несколько научных произведений, нередко число их превышает десятки и сотни наименований.

Таковы наиболее важные, на наш взгляд, свойства научной информации, они оказывают непосредственное воздействие на развитие научно-информационной деятельности и являются объектом изучения многих научных дисциплин, например, информатики, документалистики, философии, теории информационных систем.

Для построения информационной системы необходима качественная оценки информации, т.е. она должна соответствовать некоторым требованиям:

- достоверность данных, т.е. степень их соответствия описываемому объекту. Достоверность обеспечивается некоторыми приемами, например, многократностью наблюдений, независимостью источников данных, проведением контроля, проверочными расчетами, экспериментами. При этом достоверность информации зависит не только от достоверности исходных данных, но и от способов их передачи и усвоения. При передаче данных по каналу связи можно наблюдать физический шум (вызванный самим каналом связи) и семантический шум (он связан со способностью получателя понять, усвоить данные). Таким образом, информация достоверна, если достоверны базовые данные и устранены помехи при их передаче и усвоении;

- полнота информации, т.е. достаточность для принятия рационального, наиболее правильного решения. При этом объем информации, особенно первичной, должен быть минимальным. Необходимо соблюдать принцип: минимум исходных данных, минимум первичной информации и максимум производной вторичной информации. Нарушение данного принципа в практических условиях приводит к избыточности данных и вызывает излишние трудозатраты и помехи при передаче и усвоении базовых данных;

- своевременность и точность информации. Процесс управления, в частности принятие и реализация решения, основывается на полученной, переработанной и усвоенной информации. Поэтому к информации предъявляются определенные требования относительно ее возраста. Информация должна быть своевременной, оперативной и не должна быть морально устаревшей, так как в противном случае она не будет содержать необходимые сведения;



– полезность информации. Вся информация, понятая получателем, является потенциально полезной. Ценность ее увеличивается в той мере, в какой она будет использована: во-первых, для обогащения знаний получателя информации о системе; во-вторых, для постановки и решения задач управления;

– избирательность информации. Для повышения ценности информации необходимо повысить ее избирательность, т.е. соответствие определенному потребителю.

## **Глава 2. Научно-информационная деятельность**

Развиваясь на протяжении тысячелетий, наука как вид деятельности закономерно пришла к необходимости общественного разделения труда, без которого невозможно сохранить темпы развития и эффективность научных исследований. В результате разделения труда в науке в качестве самостоятельных видов выделились научно-исследовательская, научно-информационная и научно-организационная работа. Каждая из этих разновидностей научного труда имеет свои задачи, цели и теоретическую основу функционирования. Задача научных исследований – это познание закономерностей окружающего мира, природы, мышления. Научно-организационная деятельность ставит своей задачей подготовку научных кадров, повышение их квалификации, обеспечение необходимых организационных, технических условий для научно-исследовательской работы. Научно-информационная деятельность решает свои задачи, которые нацелены на повышение эффективности научных исследований в первую очередь за счет информационного обеспечения потребностей ученых.

Таким образом, научно-информационную деятельность (НИД) можно определить как организационно оформленную разновидность научного труда, которая выполняется в целях повышения эффективности исследований и заключается в сборе, аналитико-синтетической переработке, хранении и поиске закреплённой в документах научной информации. Под аналитико-синтетической переработкой научной информации здесь понимаются ее систематизация, анализ и синтез, не доходящие до получения нового знания, что является уже собственно исследовательской задачей.

При обсуждении роли и места научно-исследовательской деятельности в современной науке можно выделить две точки зрения:

1. НИД – это разновидность технического обслуживания исследователей в обеспечении необходимой для научной работы

информацией, но к науке она прямого отношения не имеет. Эта точка зрения смыкается с мнением некоторых библиографов и библиотекovedов, которые рассматривают научно-информационную деятельность как разновидность библиотечно-библиографического обслуживания;

2. НИД приравняется к научным исследованиям. Сторонниками такой точки зрения выступают некоторые специалисты по информатике и, особенно, работники органов научно-технической информации.

Вероятно, как всякое категоричное крайнее мнение, обе точки зрения не совсем верны, поскольку в них роль и значение НИД либо слишком принижаются, либо толкуются излишне широко. Несомненно, что научно-информационная деятельность является необходимой составной частью научного труда и представляет собой как бы первый подготовительный этап любого исследования. Он связан с уточнением научной проблематики путем освоения результатов и достижений предшественников и современников, работающих в данной отрасли науки, сбором, систематизацией и обработкой необходимых для исследования сведений и данных. В частности, на этом этапе много времени и сил отнимают такие, в целом чисто технические операции, как подбор литературы, научных документов, относящихся к предмету исследования. В исторической науке необходимы тщательная работа с опубликованными и архивными источниками, их поиск и анализ. В течение столетий эта работа выполнялась учеными при помощи библиотек и архивов. В результате исследователи вынуждены были тратить 70—80% своего времени на поиск информации.

Первые и наиболее важные формы и методы научно-информационной деятельности стали складываться в рамках библиотечно-библиографического обслуживания. Библиотеки вплоть до середины XX в. были, пожалуй, единственным общественным институтом, который накапливал, анализировал и обобщал опыт работы с информацией.

Бурное развитие науки в XIX-XX вв. неизбежно вело к тому, что без должной организации научно-информационной деятельности и выделения ее в самостоятельную отрасль невозможен был дальнейший прогресс. При этом нужно отметить, что некоторые виды научно-информационной деятельности, к которым относятся прежде всего реферирование, подготовка обзоров, требуют от выполняющих ее людей высокого уровня знания и применения теории и методов соответствующих наук, а кроме того, специальных навыков работы с первичной информацией, ее отбором, систематизацией,

информационным анализом. Для подготовки и привлечения специалистов такого уровня научно-информационная деятельность может и должна быть организована в рамках общественных и государственных институтов.

11 мая 1962 г. было принято постановление Совета Министров СССР «О мерах по улучшению организации научно-технической информации в стране», которое определило структуру информационной службы и установило основные принципы ее работы. Для решения задач научно-информационного обеспечения был создан ряд общесоюзных и отраслевых институтов научно-технической информации (НТИ), центров и система информационных служб.

Структура органов научно-технической информации строилась по иерархическому принципу и подразделялась на несколько уровней:

- центральные институты научно-технической информации;
- центральные отраслевые институты научно-технической информации;
- региональные (республиканские) центры информации;
- отделы (бюро) НТИ научно-исследовательских, проектно-конструкторских организаций;
- бюро технической информации предприятий.

Важным звеном в сети информационных органов являются технические, научно-технические, научные библиотеки, архивы, музеи, которые помимо своих основных задач решают некоторые вопросы научно-информационного обеспечения.

Уже в 70-е гг. сеть органов научно-технической информации насчитывала 10 всесоюзных, 82 отраслевых информационных центра, 15 республиканских институтов НТИ, 68 территориальных центров и около 9 тыс. отделов (бюро) информации<sup>1</sup>.

Основная задача деятельности органов научно-технической информации – это обработка и подготовка информационных материалов по опубликованным источникам и создание эффективных информационно-поисковых систем.

Особое место среди органов НТИ занимают центральные институты, которые к настоящему времени превратились в крупнейшие информационные центры мирового значения – Всероссийский институт научно-технической информации (ВИНТИ), Институт научной информации по общественным наукам (ИНИОН). Одновременно они являются головными, координационными научно-исследовательскими

---

<sup>1</sup> Никитин П.И. Информационно-поисковые системы: Теоретические основы информационного поиска. М., 1972. С.15.

центрами по вопросам теории и методики научно-информационной деятельности.

Эти институты выполняют следующие задачи:

- подготавливают и выпускают информационные издания по данным научно-технической литературы;
- осуществляют справочно-информационное обслуживание по разовым и постоянным запросам;
- ведут исследования по основам теории и организации научно-технической информации;
- осуществляют функции головных организаций по созданию единой государственной системы научно-технической информации;
- создают и эксплуатируют информационно-поисковые системы различных типов;
- готовят кадры для работы с научной информацией.

Отраслевые органы научно-технической информации представлены научно-исследовательскими институтами информации; центральными отраслевыми институтами научно-технической информации и др. Отраслевые органы накапливают и систематизируют информационные материалы по отдельным отраслям науки и техники, что позволяет осуществлять технико-экономические исследования.

Низовые органы научно-технической информации (отделы и бюро), как правило, функционируют в научно-исследовательских и проектно-конструкторских институтах, на предприятиях и выполняют свои задачи, содержанием которых является обобщение и распространение достижений науки, техники, а также информирование ученых, инженерно-технических работников о научных, технических и технологических новшествах. В их задачи входит поиск информации по тематике работ институтов и предприятий, ее систематизация и обобщение; подготовка информационных материалов по темам, разрабатываемым в данной организации.

Между центральными, отраслевыми, региональными и низовыми органами информации существуют тесные интеграционные связи, реализуемые через потоки информации. Центральные и отраслевые институты концентрируют свои усилия по сбору, обработке научно-технической литературы, издаваемой в нашей стране и за рубежом и по выпуску информационных изданий, которыми обеспечиваются региональные центры и низовые органы, библиотеки страны. Низовые органы имеют возможность сосредоточиться на обслуживании потребителей информации и в свою очередь пополняют информационные фонды центральных органов дополнительными

источниками информации: отчетами о научно-исследовательской работе и др.

Основным средством информационного обеспечения, которое создается в рамках органов НТИ, являются реферативные журналы; сигнальная информация, экспресс-информация; аннотированные библиографические указатели, обзоры; справочники и другие материалы.

Реферативный журнал – это специализированное справочно-библиографическое периодическое издание, целью которого является регистрация и систематизация научно-технической литературы. Реферативный журнал как особый вид периодических изданий появился еще в конце XVIII – начале XIX в. В 1769—1776 гг. в Лейпциге стал издаваться журнал «Vollständige Auszüge aus den besten chirurgischen Disputen aller Akademien» («Полные выдержки из лучших хирургических диспутов всех академий»). Он считается первым реферативным журналом в мире. Несколько позднее аналогичный журнал стал издаваться в Лондоне, а затем и в других странах. Потребность в такого рода информационных изданиях возникла в естественных науках, которые в этот период развиваются особенно бурно.

В России наиболее известным информационным изданием был «Указатель открытий по физике, химии, естественной истории и технологии», он публиковался в Санкт-Петербурге в 1824—1831 гг. Большую роль играл и другой реферативный журнал — «Медицинское обозрение», выпускавшийся с 1874 г. вплоть до 1918 г. и насчитывающий 89 томов, а также «Библиографический указатель статей железнодорожной периодической литературы» (СПб, 1833 – 1916). Кроме того, большинство научных журналов этого периода имели реферативные отделы.

Первым советским реферативным изданием был журнал «Сообщения о научно-технических работах в Республике» (1920—1932). В 20—30-е гг. издавался еще ряд реферативных сборников, таких как «Центральный медицинский (реферативный) журнал» (1928—1941); «Химический реферативный журнал» (1931—1935, 1938—1941); «Новости технической литературы (1936—1953). Однако вплоть до 50-х гг. подготовка и издание этих информационных журналов осуществлялись в рамках академических институтов, и поэтому они имели ряд существенных недостатков: нарушение периодичности, длительные сроки выхода справочников, недостаточная полнота охвата публикаций и т.д.

В 1952 г. в системе Академии наук был создан Институт научной информации, на него возложили задачу подготовки и издания реферативных журналов по естественным и точным наукам и по технике. Уже в 1953 г. силами института были подготовлены и выпущены первые реферативные журналы по астрономии, математике, механике, физике, биологии и другим наукам. Позднее этот институт был реорганизован во Всесоюзный институт научной и технической информации (ВИНИТИ) при Государственном комитете Совета Министров по науке и технике и Академии наук СССР. ВИНИТИ значительно расширил работу по подготовке реферативных журналов, охватив практически все отрасли научного знания. Для подготовки реферативных журналов и других информационных материалов ВИНИТИ использует около 17 тыс. наименований журналов, более 4 тыс. трудов и отчетов институтов, 110 тыс. патентов и много других изданий. Кроме ВИНИТИ реферативные журналы готовят и издают ИНИОН, Научно-исследовательский институт медицинской информации (по медицине), Центральный институт научной информации по строительству (по строительству и архитектуре), Центральный институт патентной информации и др.

В настоящее время реферативный журнал остается одним из наиболее распространенных информационных журналов, роль которого трудно переоценить. Он позволяет специалисту в короткие сроки ознакомиться со всеми новыми публикациями, которые появляются по интересующему его вопросу в нашей стране и за рубежом. Реферативный журнал практически по всем отраслям науки, а также по многим межотраслевым проблемам выходит ежемесячно, снабжен подробными авторскими, предметными и другими указателями, которые значительно облегчают поиск информации.

Реферативный журнал выполняет следующие функции:

- служит средством текущего оповещения ученых и специалистов обо всей публикуемой научно-технической литературе по данной отрасли (функция сигнальной информации);
- является хорошим инструментом для ретроспективного поиска научных документов по соответствующим отраслям и проблемам;
- смягчает последствия дифференциации науки и рассеяния публикаций;
- позволяет в определенной степени \*преодолеть первичный межъязыковой барьер, так как реферируются и научные публикации, вышедшие на иностранных языках. В реферативном журнале ВИНИТИ,

например, публикуются рефераты и библиографические описания, составленные на статьи, описания изобретений, сборники, монографии, выходящие более чем в 130 странах мира на 66 языках;

- помогает ученому или специалисту следить за достижениями в смежных областях науки и использовать эти достижения в своей отрасли;

- служит средством косвенной оценки ценности научной публикации, поскольку издания, не содержащие ценной информации, либо не отражаются в реферативном журнале, либо представлены лишь в виде аннотации или библиографического описания.

Реферативный журнал был и остается одним из наиболее значимых инструментов информационного поиска. По данным социологических исследований журналом пользуются 80,4% всех опрошенных научных работников, 78% аспирантов, 65,3% работников информационных служб, 57,8% преподавателей высших учебных заведений. Интересно, что наиболее высок спрос на реферативные журналы по биологии (около 88,7% всех опрошенных специалистов), химии, геологии, физико-математическим наукам. Представители общественных, гуманитарных наук в меньшей степени имеют тенденцию к использованию реферативных журналов, возможно, потому что темпы роста этих наук и приращения знаний здесь ниже.

Из других видов информационных изданий большой интерес представляет сигнальная информация (СИ). Бюллетени сигнальной информации служат для быстрого оповещения ученых и специалистов о всех новых публикациях по интересующим их отраслям, предметам, проблемам. Издания СИ являются важным дополнением к реферативным журналам, они ориентированы на охват публикаций по какому-либо одному предмету или проблеме и предназначены для достаточно узкой группы специалистов, насчитывающей, по крайней мере, несколько сотен человек.

К изданиям СИ предъявляются следующие требования:

- время оповещения о новых публикациях не должно превышать 1—2 недель с момента получения их оригиналов;

- по возможности полно охватывать все публикации по данной проблеме;

- издания должны иметь соответствующий поисковый аппарат, позволяющий легко ориентироваться и вести поиск. Поэтому библиографическое описание публикации начинается не с имени автора, а с названия работы.

Выделяются следующие типы изданий СИ:

- бюллетени, составляемые из оглавлений выпускаемых в свет важнейших научно-технических журналов, обязательно снабжаемые авторскими указателями;
- библиографические указатели, в которых под предметными заголовками приводятся полные библиографические описания соответствующих публикаций.

ВИНИТИ с 1966 г. приступил к изданию СИ и уже в 1976 г. публиковал около 100 названий бюллетеней и указателей — по 24 номера ежегодно.

Помимо подготовки и публикации информационных изданий органы научно-технической информации осуществляют и другие виды информационного обслуживания, среди которых следует выделить избирательное распространение информации (ИРИ), а также ретроспективный поиск документов и информации.

Избирательное распространение информации как вид информационного обслуживания обеспечивает эффективное удовлетворение индивидуальных информационных потребностей ученых и специалистов. Эта форма обслуживания издавна практиковалась библиотеками, архивами. Однако без современной техники и использования компьютеров решение задачи регулярного и оперативного снабжения потребителей необходимыми документами невозможно. Термин «избирательное распространение информации» был введен в обращение Х.П.Луном, который в 1958 г. предложил создать систему, основанную на использовании ЭВМ. Идеи Луна использовались для разработки таких систем в ряде американских компаний в 60-е гг., и они прочно вошли в практику информационной работы.

Под информационными системами ИРИ понимаются такие системы, которые обеспечивают срочное оповещение потребителей информации о публикациях, представляющих для них потенциальный интерес. Потребитель информации заранее сообщает информационному органу перечень своих потребностей, выраженных в поисковых признаках и характеристиках (фамилии авторов, предметные рубрики, ключевые слова и т.д.). Информация, поступающая в справочно-информационные фонды, анализируется по этим признакам, и в случае их совпадения заказчику высылаются интересующие его информационные материалы.

В СССР практическая эксплуатация систем избирательного распространения информации началась в 1964 г., и сейчас действуют ее



автоматизированные варианты. Современные автоматизированные системы ИРИ функционируют следующим образом: поступающие в систему документы индексируются (иногда аннотируются и реферируются) специалистами. На каждый вводимый в систему документ составляется запись по определенной форме, пригодной для ввода в компьютер: библиографическое описание документа, поисковый образ (ключевые слова), возможно, текст аннотации или реферата. С определенной периодичностью, например, ежедневно, эти записи вводятся в компьютер, где уже находятся поисковые профили абонентов, представленные поисковыми предписаниями. При вводе информации поисковые образы документов автоматически сравниваются с профилями абонентов. Если сравнение указывает на соответствие, то ЭВМ печатает имя абонента, библиографическое описание документа, его аннотацию или реферат. Эти распечатки направляются абоненту, который по своему выбору может запросить полную копию документа.

Ретроспективный поиск информации связан с выполнением разовых запросов по какой-то тематике. Это достаточно сложный и дорогостоящий вид информационного обслуживания, но в условиях развития рыночных отношений он имеет большие перспективы.

В последнее время все большее распространение получают системы ИРИ, основанные на использовании информационно-поисковых массивов, которые подготавливаются в централизованном порядке крупными информационными центрами и институтами и распространяются по подписке низовым информационным органам. Как считают специалисты, оптимальная схема информационного обслуживания должна содержать два уровня: 1) централизованная обработка научно-технических изданий с целью подготовки бюллетеней сигнальной информации, которые распространяются по подписке; 2) избирательное распространение информации среди ученых и специалистов, а также решение задач ретроспективного поиска, которые осуществляются силами низовых органов НТИ. Такая схема позволяет крупным информационным центрам сосредоточить все усилия на аналитико-синтетической обработке научно-технической литературы и создании центрального информационно-справочного фонда. А низовые информационные органы, получая такие фонды в готовом виде, могут сконцентрировать свое внимание на обслуживании потребителей.

Подготовка информационных изданий требует много труда, средств и знаний. К созданию реферативных журналов и сигнальной информации привлечены десятки тысяч людей. Только в подготовке

реферативных журналов ВИНТИ принимают участие более 30 тыс. штатных и внештатных сотрудников. Сбор, систематизация, оценка, обзор научных документов – это сложная задача, решение которой требует знаний и профессиональных навыков, связанных с освоением видов аналитико-синтетической переработки информации.

Главная задача аналитико-синтетической переработки информации, содержащейся в каком-либо множестве документов, заключается в группировке, сопоставлении, оценке и обобщении ее, а также в представлении обобщенной информации в более компактном виде, удобном для ее практического применения.

В научно-информационной деятельности различают следующие виды аналитико-синтетической обработки научных документов и информации:

- библиографическое описание;
- индексирование;
- аннотирование и реферирование;
- перевод научных документов;
- сбор, критическая оценка, систематизация и обобщение научно-технических данных;
- составление обзоров научно-технической литературы.

Эти виды переработки научной информации и документов перечислены в порядке возрастания сложности и затрат умственного труда, которые необходимы для их осуществления.

Библиографическое описание представляет собой процесс и результат составления по определенным правилам перечня сведений о каком-либо документе, что позволяет найти документ среди других. Главный объект анализа в данном случае – это не содержание документа и не сам текст, а его внешние опознавательные признаки. По разработанным в течение веков правилам определяются заглавие документа, имя автора, надзаголовочные и подзаголовочные данные, выходные данные и другие опознавательные признаки. В подзаголовочных данных обычно указываются: целевое назначение документа (учебное пособие, справочник и т.п.); форма документа (альбом чертежей и т.п.); имена лиц, принимавших участие в подготовке документа (переводчик, редактор, иллюстратор и т.д.); сведения о повторности издания. Выходные данные отражают сведения о издании: место издания, год, имя издателя и т.п. Кроме того, в библиографическом описании приводятся сведения об объеме документа (общее число или номера страниц), о количестве содержащихся в нем иллюстраций, таблиц и т.д.

В библиографическом описании все элементы следуют строго один за другим: имя автора, заглавие, потом подзаголовочные данные, выходные данные, количественная характеристика, надзаголовочные данные, примечания. Построенное таким образом библиографическое описание позволяет однозначно идентифицировать каждый научный документ, без чего невозможно создание информационных систем, таких, как, например, библиотечные каталоги, реферативный журнал, библиографические указатели.

В библиотечно-библиографической работе необходимы полные библиографические описания, однако для нужд НИД отдельные элементы его могут быть опущены как избыточные. Поэтому в реферативных журналах, в бюллетенях сигнальной информации используется сокращенный вариант библиографического описания (без некоторых видов подзаголовочных данных, примечаний и т.д.), а также может быть изменен порядок построения элементов. В частности, при поиске необходимых документов исследователи обращают основное внимание на заглавие, так как именно оно позволяет достаточно точно судить о содержании документа. Поэтому в реферативных журналах и бюллетенях сигнальной информации библиографические описания начинаются не с имени автора, а с заглавия, имя автора приводится после заглавия и подзаголовочных данных.

Более сложными видами информационного анализа и синтеза являются индексирование, аннотирование и реферирование.

Индексирование – это процесс выражения центральной темы или содержания документа в терминах информационно-поискового языка, т.е. создание поискового образа документа. Для этого могут быть использованы индексы классификации, предметные заголовки или ключевые слова.

Аннотацией называется связный текст, который кратко выражает центральную тему или предмет исследования какого-либо документа. Объем аннотации обычно не превышает 150—200 знаков.

Рефератом называется связный текст, в котором отражены центральная тема или предмет исследования какого-либо документа, а также цель, используемые методы и основные результаты исследования. Средний объем реферата обычно составляет 1000-1500 знаков.

При индексировании, аннотировании и реферировании главным объектом анализа является уже содержание документа, его текст. В ходе анализа необходимо установить: центральную тему, или объект исследования, а также основные идеи и факты, связанные с этой темой, или объектом; соответствие данного документа научным и

практическим интересам группы ученых или специалистов. Для выполнения данной работы требуется, по крайней мере, знание основ той отрасли науки, к которой относится аннотируемый или реферируемый документ, иначе трудно отделить главное от второстепенного, оценить результаты исследования. Таким образом, указанные виды аналитико-синтетической обработки информации должны осуществлять специалисты.

Еще более сложным и трудоемким видом аналитико-синтетической переработки научной информации является перевод документов с одного естественного языка на другой. Такой перевод должен обеспечить точную передачу на другом языке научной информации, передача особенностей авторского стиля оригинала здесь не является главной задачей в отличие от переводов художественных произведений. Сложность при переводе документов заключается в том, чтобы наиболее полно передать смысл документа, логику развития мысли, систему доказательств. Вместе с тем многие лингвисты вообще ставят под сомнение возможность адекватного перевода любого текста с одного языка на другой, что обостряет проблему искажения информации. Изучением внутренних механизмов такого процесса занимается наука о переводе.

Особыми видами аналитико-синтетической переработки научно-технической литературы и информации являются сбор, систематизация, критическая оценка и обобщение научно-технических данных. В научной литературе, периодических изданиях, справочниках в настоящее время публикуется огромное количество показателей, результатов наблюдений, конкретных научно-технических данных. В науке и технике производится множество разнообразных измерений, но на практике оказывается, что наши знания о явлениях скудны и недостаточны. Здесь мы вновь сталкиваемся с проблемой рассеяния информации. Поэтому возникает потребность в сборе, систематизации всех данных о явлениях и процессах, чтобы сделать их доступными для всех ученых. С другой стороны, достаточно часто возникает ситуация, когда аналогичные показатели в разных документах расходятся в своих значениях. Необходимо выяснить причины этих расхождений, устранить возникающие неопределенности, установить истинное значение показателя. В рамках данного вида аналитико-синтетической переработки информации проводятся систематизация аналогичных научно-технических данных, их фильтрация, вырабатываются рекомендации, какие значения данных следует считать наиболее предпочтительными и т.д.

Самым сложным видом аналитико-синтетической переработки научной информации является составление обзоров научно-технической литературы. Составителями обзоров могут быть лишь высококвалифицированные специалисты в соответствующих отраслях науки и техники, обладающие обширной научной эрудицией и большим практическим опытом.

Научные обзоры играют исключительно важную роль в современной научной деятельности. Обычно ученый или специалист справляется с потоком появившихся научных документов в своей области, пока мощность этого потока не достигает некоторой критической величины. Когда такой предел достигнут и ресурсы человеческого организма исчерпаны, ученый отказывается от попыток в полной мере использовать литературу и ищет ей замену. Единственный способ решить радикально эту задачу заключается в организации систематического сбора, анализа и оценки всех появляющихся научных документов по основным отраслям, предметам и проблемам, а также в изложении полученной информации в компактном и удобном для восприятия виде. Иначе говоря, речь идет о подготовке обзоров.

Обзором называется синтезированный текст, в котором дается сводная характеристика какого-либо вопроса или ряда вопросов, основанная на использовании информации, извлеченной из некоторого множества отобранных для этой цели первичных документов за определенное время. Обзор не обязательно должен представлять собой самостоятельный документ, он может быть лишь частью документа.

Различают три основных вида обзоров: библиографические, реферативные, аналитические.

Библиографический обзор – это обзор, содержащий характеристику первичных документов как источников информации, появившихся за определенное время или объединенных по какому-либо общему признаку. Подготовка библиографических обзоров осуществляется главным образом в библиотеках.

Реферативный обзор – это обзор, содержащий сводную характеристику вопросов, рассмотренных в исходных первичных документах, но не дающий критической оценки приведенной в этих документах информации.

Аналитический обзор (критический) – это обзор, содержащий всесторонний анализ всей приведенной в первичных документах информации, ее аргументированную оценку и обоснованные рекомендации по использованию. В критическом обзоре составитель оценивает информацию исходных документов и на основе логического

анализа высказывает свои собственные суждения. Нужно отметить, что обзоры стареют значительно медленнее, чем первичные научные документы.

Обзоры чаще всего публикуются в виде журнальных статей и только в некоторых отраслях науки (например, океанографии) – в виде монографий. В периодических изданиях печатается около 67% всех обзоров, 23% – в продолжающихся изданиях типа «Ежегодный обзор...», «Итоги науки и техники» и всего 10% – в виде монографий. Среди отечественных обзорных изданий выделяются выпуски типа «Итоги науки ...», ежегодник «Наука и человечество», который издается с 1963 г. Средний объем обзорной статьи составляет около 20 стр.

Возникает проблема, как выделить среди научных изданий обзорные, их соотношение, по мнению американских ученых, составляет примерно 1:50, т.е. на 50 научных публикаций приходится одна обзорная. Исследования А.М.Вудведа показали, что документ является обзорным, если он отвечает следующим требованиям:

- содержит более 100 библиографических ссылок;
- в его заглавии содержатся ключевые слова «обзор», «достижения», «итоги», «результаты» и т.п.; из заглавия ясно, что документ является обзорным;
- он опубликован в каком-либо обзорном периодическом или продолжающемся издании.

Документ может быть обзорным, если:

- содержит от 40 до 100 библиографических ссылок (чем больше ссылок, тем выше вероятность того, что он является обзорным);
- отмечен или индексирован как обзор;
- имеет одного автора;
- из его заглавия следует, что это обзор.

Из-за быстрого роста количества выходящих первичных документов потребность в обзорах резко возросла, поэтому необходима организованная система подготовки обзорных научных документов. Главная ответственность за подготовку обзорной информации ложится на учреждения, осуществляющие различные процессы научной коммуникации, т.е. на органы научной информации и научные издательства. Они определяют тематику обзоров и периодичность их подготовки, работают с авторами; подбирают научные документы по тематике составляемых обзоров; оказывают помощь авторам в редактировании и техническом оформлении рукописей. В настоящее время почти 2/3 всех обзоров подготавливаются вне системы научно-технической информации. Это означает, что подготовка обзоров — во

многом стихийно развивающийся процесс. Только два обзорных издания выпускаются информационными органами, поэтому они отличаются единством методики подготовки и регулярностью опубликования.

В течение последних десятилетий в системе научно-технической информации приобретают все большее значение так называемые центры анализа информации. Одной из главных задач таких центров является подготовка аналитических обзоров.

## **Глава 3. Информационные системы**

### **3.1. Виды информационных систем**

Информационные системы в управлении или в любом другом виде деятельности – науке, просвещении, производстве — имеют общие закономерности функционирования, которые изучает целая совокупность дисциплин, в частности таких, как информатика, теория информационных систем и информационного поиска, документалистика и др.

Основное предназначение любой информационной системы – это удовлетворение информационных потребностей пользователей системы, а выполнение данной функции в значительной мере связано с поиском информации. Это одно из основных понятий теории информационных систем. Информационный поиск связан с нахождением документов или фактов, обобщением каких-либо данных или извлечением информации из многочисленных источников. Исходя из этого, под информационным поиском понимают совокупность ряда логических и технических операций, конечной целью которых является извлечение по заданным признакам всех документов или фактографических данных, находящихся в информационном массиве и отвечающих на поставленный вопрос.

С понятием информационного поиска тесно связано и определение информационной системы. ИС — это системы, в которых осуществляются информационные процессы, составляющие полный цикл обращения информации (регистрация, формирование баз данных, хранение, обработка, поиск, передача, представление в удобной для потребителя информации форме). Разновидностью информационных систем являются информационно-поисковые системы.

Структурно информационные системы представляют собой совокупность информационного массива (баз данных), информационно-поискового языка, правил обработки, поиска и выдачи информации, а

также технических средств, с помощью которых осуществляется процесс хранения, регистрации, поиска, выдачи информационных материалов.

Изучение информационных систем, их структуры и закономерностей функционирования началось в первую очередь с информационно-поисковых систем, которые получили большое распространение в научно-информационной деятельности, управлении, библиографии и архивном деле. Долгое время считали, что предметом теории информационно-поисковых систем (ИПС) является изучение только технических аспектов поиска: перфокарт, микрофильмовых селекторов, компьютеров и других средств, используемых при поиске. Однако развитие теории и практики применения ИПС показало, что в основе ручных, механизированных и автоматизированных ИПС лежат общие принципы и закономерности.

Понятие ИС включает все виды систем, начиная с традиционных библиотечных каталогов и картотек и кончая механизированными, автоматизированными системами. Если историю ИС начинать со времени возникновения первых картотечных и каталожных систем, то она уходит в глубь веков. Один из первых каталогов древности обнаружен в древнем Шумере – это список литературы на глиняной табличке, относящийся к 2000 г. до н.э. На протяжении последующих 4 тыс. лет происходит постепенное проникновение ИС в различные сферы деятельности – науку, научно-информационную деятельность. Картотеки, каталоги, справочники, указатели – это наиболее распространенные разновидности информационных систем, с которыми мы сталкиваемся и в обыденной жизни, и в профессиональной деятельности.

Эти системы называются часто *ручными*, поскольку они предполагают ручные приемы работы пользователей системы с ее информацией. Носителем информации в ручной системе чаще всего являются библиографическая и каталожная карточки, которые в определенном порядке располагаются в отделениях (ящиках) картотеки. Иногда в таких системах используются перфокарты. К ручным системам относятся также указатели, справочники, изданные в виде книги, журнала или другого печатного издания. Среди всех ИПС они наиболее распространены и широко доступны, удовлетворяя наши повседневные и профессиональные потребности в информации. Недостатком таких систем является то, что они требуют много труда при их создании и относительно много времени – для поиска нужной информации.



Проблема сокращения времени поиска, повышения его точности и полноты способствовала появлению в 30-е гг. XX в. механизированных систем, в которых стали применять специальные технические приспособления, ускорявшие поиск информации. Это могли быть специальные спицы, с помощью которых проводили операции сортировки и группировки информации, причем скорость поиска информации возрастала в 10 и более раз. В это время появились машины для перфорации, считывания перфокарт, селекторы для работы с микрофильмами. Основными носителями информации в механизированных ИПС стали перфокарты и микрофильмы.

Машинные перфокарты были изобретены и впервые использованы для механизации различных статистических работ еще в конце XIX в. Первая перфокарта с краевой перфорацией была изобретена еще в 1904 г. В 1915 г. ученый Г.Тейлор получил патент на способ поиска информации с помощью суперпозиционных перфокарт. Но практически эти изобретения стали применяться для создания ИПС лишь в 1939 г. В это же время была применена для поиска информации и микрофильмовая техника. Механизированные ИПС не получили такого широкого распространения, как ручные. Чаще всего они создавались и использовались в статистике, патентных информационных системах, частично – в управленческих документных системах.

Конец 50-х – начало 60-х гг. были ознаменованы появлением автоматизированных информационных систем (АИС) с использованием компьютерной техники и новых носителей информации, таких, как магнитные ленты и диски, лазерные и оптические диски. В автоматизированных системах поиск информации происходит в считанные секунды. АИС в настоящее время развиваются особенно ускоренными темпами, проникая во все сферы деятельности человека. С появлением автоматизированных систем начался совершенно новый, качественно отличный этап развития информационных технологий, призванный преодолеть информационный кризис.

Задачи, решаемые с помощью информационных систем, позволяют выделить основные направления автоматизации – производство, управление, наука, проектирование и др.

Автоматизированная информационная технология должна решать следующие задачи:

- приблизить пользователя к вычислительным и информационным ресурсам;

– сократить время обращения информации от ее создания до потребления.

Компоненты новой информационной технологии: а) информационно-вычислительные сети (ИВС); б) банки данных (БНД).

Информационные сети представляют собой совокупность ЭВМ, средств коммуникаций, аппаратуры для передачи данных, программных средств и подразделяются по масштабам (на глобальные – международные, региональные, локальные), по топологии (централизованные, децентрализованные, открытые, закрытые).

Банк данных – это информационная система, предназначенная для поддержания динамической модели предметной области с целью обеспечения информационных потребностей пользователя.

Под информационной технологией банков данных мы будем понимать совокупность знаний о методах и приемах функционирования банков данных, его компонентах и способах их объединения в систему. Банк данных состоит из следующих компонентов: технических, информационных, лингвистических, программных, организационных, т.е. в данном перечне мы опять же повторяем все составные элементы информационных систем, которые рассматривали ранее.

Технический комплекс представляет собой совокупность средств ВТ, кроме ЭВМ он включает запоминающие устройства и устройства ввода—вывода.

Информационный комплекс состоит из совокупности информационных моделей и баз данных, схем информационных моделей различного уровня абстрагирования, словаря данных.

Лингвистический комплекс включает языки программирования, языки описания, языки манипулирования данными.

Программный комплекс – системы управления базами данных (СУБД), операционные системы и прикладные программы. СУБД обеспечивают управление базами данных и доступ к данным. Прикладные программы – удовлетворение конкретных информационных потребностей пользователей. Операционные системы предназначены для организации всего вычислительного процесса в ЭВМ.

Организационный комплекс составляют пользователи, администратор банка данных и его группа, специалисты по эксплуатации технического и программного комплекса.

Таким образом, к середине XX в. сформировались три основные типа ИС: ручные, механизированные и автоматизированные,

отличающиеся уровнем технической оснащенности, носителями информации и быстродействием.

Помимо этой классификации в теории информационных систем используется еще одна классификация ИС в зависимости от характера заложенной в них информации, а также типа информационных потребностей, которые они удовлетворяют.

Существуют два типа информационных потребностей: 1) в сведениях об источнике информации, о документах, их адресах; 2) в конкретных данных, сведениях, фактах, т.е. так называемые фактографические потребности. В соответствии с этим различаются:

- документальные системы, нацеленные на поиск документов, адресов и описаний документов;
- фактографические системы, предназначенные для поиска конкретных фактов.

Задача документальных информационно-поисковых систем (ДИПС) – обработка документов, нахождение их по тем или иным признакам, а также выдача документов, их копий и рефератов. Элементом описания в документальных системах является документ, его внешние опознавательные признаки, адрес местонахождения. Под документом мы понимаем информацию, закрепленную определенным образом на материальном носителе. Поэтому документом считаются и управленческая документация, и статьи, монографии, а также любые другие издания, исторические источники и т.д. в рукописном, печатном или машиночитаемом виде. С документальными системами все хорошо знакомы. Работая в библиотеке, мы в первую очередь обращаемся к библиотечному каталогу, часто пользуемся библиографическими указателями, реферативными изданиями и другими справочниками. В архиве для поиска информации созданы путеводители по фондам, описи фондов, каталоги документов – все это системы документального типа, поскольку в результате поиска мы получаем адреса документов, их опознавательные признаки и описания. В управлении традиционно широко используются самые разнообразные виды документальных систем: номенклатура дел, каталоги для организации контроля исполнения документов, журналы регистрации входящей и исходящей документации.

Фактографические информационно-поисковые системы (ФИПС) предназначены для анализа, поиска и выдачи сведений или фактов, составляющих содержание документа. Фактографические системы также достаточно широко распространены – это, например, телефонные и адресные справочники. В каждом киоске сейчас можно купить газеты

с информацией о предложении товаров, услуг, вариантов обмена. Они представляют собой фактографические информационные системы, нацеленные на удовлетворение наших конкретных потребностей. В бизнесе актуальны системы, позволяющие получить определенную коммерческую информацию о потребителях товаров и конкурентах – это всякого рода коммерческие справочники, где можно найти адреса, телефоны фирм, перечень товаров и услуг, которые они предоставляют. Еще в начале века в Перми в течение 15 лет издавался подобный Адрес-календарь, включавший тысячи описаний торговых фирм и компаний, действовавших на уральском рынке. Для историков также необходимы свои фактографические системы – это, например, хронологические таблицы, именные указатели и т. д. В управлении к разряду фактографических систем информации можно отнести системы технической документации, кадровой, бухгалтерской и финансово-учетной документации и т.д.

Несмотря на знакомые нам названия, фактографические ИС все же менее распространены, чем документальные, хотя потребности в них растут очень быстро. Это связано с двумя моментами: с одной стороны, в настоящее время накоплены огромные массивы данных, эффективное использование которых невозможно без автоматизированной обработки и поиска. А с другой стороны – в обществе отсутствует исторически сложившаяся социальная система, организация, которая бы целенаправленно осуществляла сбор, обработку, хранение и поиск фактографических данных. Если для разработки документальных систем в обществе создаются специальные государственные структуры – это органы научно-информационной деятельности: институты информации ВИНТИ и ИНИОН, региональные информационные центры, отраслевые институты информации, библиотеки, музеи и архивы, в их задачи входит создание, совершенствование документальных ИПС, они также занимаются обобщением опыта создания таких систем, теоретическими проблемами информационного поиска.

Фактографические информационные системы создаются частными усилиями организаций, учреждений, научных и коммерческих. Эти усилия, как правило, разрозненны, отсутствуют единые принципы разработки таких систем, и далеко не всегда они способны достаточно полно удовлетворить информационные потребности.

Уточним понятие «факт», чтобы яснее представить возможности фактографических ИПС. Обычно фактом называют конкретное событие, установленное с полной достоверностью, имена и даты. Факт должен

быть достоверным и надлежащим образом описанным с указанием, где, когда, кем он наблюдался и был зарегистрирован.

Факты и данные разбиваются на следующие разновидности:

- результаты непосредственных наблюдений, измерений;
- имена собственные, т.е. географические, фирменные и прочие наименования;
- фотографические изображения;
- хронологические характеристики явлений и процессов;
- характеристики (признаки) предметов, явлений и процессов;
- чертежи, графики и т.д.;
- географические карты и схемы;
- математические модели и формулы.

Документальные и фактографические ИС отличаются друг от друга прежде всего содержанием хранимых в них материалов, т.е. содержанием информационных массивов, а также информационно-поисковыми языками. При документальном информационном поиске документы разыскиваются потому, что ученый или специалист надеется найти в них необходимые ему сведения и факты. Следовательно, конечной целью документальной ИС является также поиск фактов. Это позволяет сделать вывод об отсутствии принципиальных различий между фактографическими и документальными системами в теории, методике и практике их построения и использования. Отмеченные отличия в основном относятся к сфере ИПЯ, который должен строиться с учетом специфики описываемых данных.

Отсутствие существенных различий между ФИПС и ДИПС позволяет разработать универсальные методики и соответственно пакеты программ, предназначенные для поиска на ЭВМ документов и фактов. Примером такого универсального пакета программ может служить STAIRS (система хранения и поиска информации), разработанная в 1971 г.

### **3.2. Информационный массив**

Рассмотрим более подробно структуру ИС. Одним из основных элементов информационной системы является информационный массив, или, как принято называть его в автоматизированных системах, база данных.

Понятие «информационный массив» имеет двоякое толкование: как массив документов и как массив записей однородных данных. В первом случае под информационным массивом понимается упорядоченный комплекс документов (это может быть коллекция печатных произведений или совокупность документов, собранных в

архиве, управленческая документация, сформированная в дела). Сюда мы также отнесем и совокупность различных документов, собранных в единую систему и предназначенную для справочно-информационной работы, картотеку, каталог и т.п.

Во втором случае – записи на магнитных носителях, обладающие общими структурными элементами и языковыми признаками.

Таким образом, информационный массив представляет собой комплекс документов или набор записей, обладающих общими структурными элементами и языковыми признаками. В документальных ИПС массив состоит из описаний источников информации или адресов документов, а в фактографических системах – из описаний фактов.

Структура информационного массива характеризуется многослойностью. Информационный массив, или база данных, состоит из элементов информации или первичных однородных объектов описания, которые в свою очередь представляют собой определенную совокупность характеристик (признаков), описывающих объект. Элементом информации может быть документ, описание документа, реферат, аннотация, факт. Характеристика – это свойство документа или факта, его поисковый признак, дескриптор, ключевое слово и т.д.

Организация массива, т.е. приведение его в определенный порядок, имеет важное значение и влияет на эффективность работы ИС. Два фактора влияют на организацию массива : 1) логика поиска; 2) носители информации, на которых формируются массивы.

Существует ряд методов организации массивов:

- последовательный метод характеризуется логическим расположением записей, где запись  $n+1$  следует за  $n$ -й записью. Такой способ используется во многих системах (например, регистрация делопроизводственной документации), он экономичен. Однако этот метод имеет существенный недостаток – затруднено изменение структуры массива. Например, включение новой записи между двумя рядом стоящими требует смещения всех записей, входящих в массив;

- цепной метод, при котором каждая запись содержит информацию, т.е. адрес, где расположена следующая запись. В каталогах цепной метод реализуется при помощи дополнительных ссылок;

- метод организации типа дерева. Массив формируется исходя из определенных принципов систематизации (например, алфавитный), при этом элемент информационного массива может иметь ссылки на массивы, содержащие дополнительную информацию. Такой метод

позволяет экономить память и время на запись повторяющейся информации.

Информационный массив обязательно подвергается обработке. Под обработкой понимается процесс его преобразования. Оно возможно путем изменения каждой отдельной записи или перестройки последовательности размещения записей. Важнейший этап преобразования массива – его упорядочение. Упорядочение – это процесс размещения в определенном порядке записей в массиве относительно друг друга. Упорядочение может выполняться путем сортировки, т.е. путем последовательного расположения всех элементов массива в зависимости от некоторых признаков (в алфавитном порядке, по индексам УДК и др.).

Прежде чем приступить к упорядочению массива, необходимо разработать его схему. Массив считается упорядоченным в том случае, если каждая запись в нем расположена в последовательности, предусмотренной этой схемой. Все характеристики в схеме располагаются по степени старшинства (например, в алфавитном порядке).

Ввод новой информации в уже созданный массив влечет за собой слияние новых записей с существующими, т.е. получение нового массива. Чтобы исключить устаревшую или ненужную информацию, предварительно необходимо осуществить поиск нежелательных записей. Операцию исключения можно рассматривать как процесс пополнения, т.е. слияния с отрицательным знаком. В результате выполнения операции «отрицательного слияния» будет получен обновленный, уменьшенный массив.

### **3.3. Технические носители информации**

Носители информации играют важную роль в комплектации и организации информационных массивов. Носителем информации называется физическая среда, используемая для записи и накопления информации. Носители информации подразделяются на дискретные (библиографические карточки, перфокарты) и непрерывные (магнитные ленты и диски). Помимо этого используется и другая классификация носителей – они делятся на ручные, или традиционные, механизированные и автоматизированные. Традиционными носителями информации являются библиографические карточки, карточки «Униформ». К механизированным относятся перфокарты, перфоленты и микрофильмы. Магнитные носители используются в ЭВМ. Кроме того, выделяется еще одна группа носителей информации – фото- и видеоносители, они могут быть использованы в ИС любого типа.

Преимущество последних носителей состоит в том, что фотоносители могут хранить не только поисковые характеристики документов, но и тексты самих документов, и изображения.

Важнейшее свойство носителей информации – информационная емкость. Последняя зависит от размера носителя, плотности записи информации и других качеств. Информационная емкость – это максимальное количество информации (информационных или поисковых признаков), которое можно хранить на носителе. Под плотностью записи информации понимается максимальное количество информации, которое может быть записано на единицу его площади.

Магнитные носители для автоматизированных систем за последние десятилетия существенно изменились. На смену магнитным лентам, применяемым в больших ЭВМ, приходят носители, используемые на персональных компьютерах, хотя в крупных банках данных информация до сих пор хранится и поддерживается на магнитных лентах. Магнитные диски в ПК подразделяются на жесткие диски (Hard Disk) и гибкие диски (Floppy Disk).

Жестким диском называют носитель информации в виде магнитного диска на твердой основе. Пользователи персональных компьютеров часто называют его сокращенным названием «винчестер», или «винт». Жесткие диски обеспечивают запись значительно больших объемов информации, чем гибкие, и для обращения к нему компьютер затрачивает существенно меньше времени. Жесткие диски встраиваются в корпус компьютера, и обычно на них хранят те программы, с которыми постоянно работает пользователь. Тип жесткого диска определяет объем хранимой информации и скорость ее записи (чтения) и зависит от типа процессора. Для AT/386 обычно отмечается емкость 120 Мбайт, для 486 – 256 Мбайт. Компьютер может комплектоваться двумя и более жесткими дисками. Емкость жесткого диска зависит от типа и класса ПК.

Кроме того, выделяют Floppy Disk (более распространенное название – гибкий диск, дискета). Дискета – это намагниченный эластичный диск из пластичного материала, предназначенный для временного хранения данных. Диск заключен в защитный кожух (бумажный или пластмассовый) с прорезью для головок чтения (записи). На практике используются дискеты 5,25 дюйма и 3,5 дюйма. 5,25-дюймовые дискеты выпускаются с различной возможной плотностью записи. Наиболее распространенные дискеты на 360 Кбайт, 720 Кбайт, 1,2 Мбайт. Для того, чтобы наглядно оценить емкость этих



носителей, отметим, что на дискете емкостью 360 Кбайт можно хранить 200 страниц текста.

В последнее время все большее распространение приобретают так называемые 3,5-дюймовые дискеты. Основное преимущество их состоит в том, что они имеют пластмассовый чехол с металлическим экраном, предохраняющий дискету от внешнего механического воздействия. Когда дискета вставляется в дисковод, экран автоматически сдвигается и открывает магнитную поверхность диска. Емкость 3,5-дюймовых дискет обычно составляет 1,4 и 0,72 Мбайта. Внешне дискеты отличаются наличием второго «окошечка»: дискеты на 720 Кбайт имеют только одно окно, а на 1,4 Мбайта – два прямоугольных выреза.

Параллельно развиваются оптические диски, на которых запись и считывание информации осуществляются лучом лазера. На такой диск можно записать около 1 млн документов по тысяче знаков, они отличаются долговечностью и надежностью. Как разновидность оптических дисков выступают видеодиски. Один видеодиск вмещает около 50 тыс. кадров цветных изображений, например, археологических находок, шедевров живописи, документальных или художественных фильмов. На видеодисках могут храниться справочные данные, тексты редких книг, журнальных статей.

### **3.4. Техническое оснащение информационных систем**

Любая информационная система, в том числе ручного типа, предполагает наличие определенного технического обеспечения. Ручные системы наиболее неприхотливы в этом отношении, для их функционирования достаточно создать нехитрое приспособление в виде стеллажа с ящиками для картотеки. Могут быть и более сложные варианты: вращающиеся, вертикальные, блочные картотеки. Разнообразные способы технической организации картотек и каталогов создаются для более удобного использования и быстрого поиска нужной информации.

Техническое обеспечение механизированных систем предполагает использование специальной техники для записи, поиска и считывания информации с перфокарт и перфолент. Для записи информации на перфокартах с краевой перфорацией, например, использовали: щипцы-компостеры для просечки перфораций; устройство для пробивки перфокарт; клавишные устройства для кодирования и др. Для поиска информации применяли различные виды селекторных устройств, например, ручные селекторные устройства в виде ящиков (с отъемными днищами) или рамок, в которых

размещается массив карт; электрические вибрационные селекторы емкостью 400 перфокарт и др. Для хранения перфокарт использовались картотечные шкафы и картотечные столы. При работе с суперпозиционными перфокартами необходимы были специальные электромеханические устройства для считывания информации.

Автоматизированные системы невозможны без применения компьютерной техники. Персональный компьютер – это небольшой комплекс взаимосвязанных устройств, каждому из которых поручена определенная функция. Каждый конкретный компьютер может работать с разным набором внешних устройств, например, с принтером, сканером и т.д. Тем не менее существует минимальная «конфигурация», т.е. минимальный набор элементов, без которых работа с ПК становится бессмысленной. Этот набор включает: 1) системный блок; 2) монитор (дисплей); 3) клавиатуру. Сердце машины и ее мозг – системный блок, с которым соединены кабелями клавиатура и монитор. Системный блок компьютера заключен в металлическую коробку (обычно серого цвета) и часто служит подставкой монитору, однако на последних моделях ПК он устанавливается отдельно.

В системном блоке размещены: процессор; один или два накопителя на жестком диске; один или два накопителя на гибком магнитном диске. Термином «процессор» обозначается следующая совокупность блоков: микропроцессор – мозг машины, который выполняет поступающие на его вход команды; оперативная память, предназначенная для временного хранения программ и данных, а также для постоянного хранения одного из блоков операционной системы; электронные схемы, управляющие элементами компьютера и обменом данных между памятью и другими устройствами запоминания и отображения информации.

Первые модели IBM PC (XT) работали на базе процессора фирмы Intel. В середине 90-х гг. в нашей стране преобладали модели IBM PC с микропроцессором 80286, 80386, 80486. Эти модели различаются быстродействием, емкостью оперативной памяти.

Монитор предназначен для вывода на экран текстовой и графической информации. В настоящее время на компьютерах наиболее широко используются цветные мониторы типа EGA, VGA, SVGA. В текстовом режиме картинки на экране этих мониторов практически не отличаются друг от друга, однако в графическом режиме качество изображения у SVGA выше, чем у VGA, а у VGA – выше, чем у EGA.

С помощью клавиатуры мы вводим алфавитно-цифровые данные и управляем работой компьютера.

Принтер (печатающее устройство) предназначен для вывода информации на бумагу. Область применения принтеров достаточно широка – от вывода обычной текстовой информации до построения сложных графических изображений (иногда – в цвете). Наибольшее распространение в работе с ПК получили принтеры трех типов: матричные, струйные и лазерные. Лазерные принтеры обладают целым рядом преимуществ перед матричными и струйными. Они дают наилучшее (близкое к типографскому) качество печати, скорость печати у них в 4-5 раз выше, и они надежны в использовании.

### **3.5. Информационно-поисковые языки**

Первой операцией при создании информационной системы является анализ содержания документов, определение формы их описания. Эффективность информационного поиска будет зависеть от качества анализа документа, информационно-поискового языка и методов индексирования.

Важнейшим элементом ИС является информационно-поисковый язык (ИПЯ). Он – ключевая проблема обработки и поиска как документов, так и фактографической информации. Важнейшим свойством информационного языка является его способность выражать необходимую информацию. Ни один естественный язык не может быть применен в качестве ИПЯ. Этому препятствуют их недостаточно четкая структура, неоднозначность (наличие омонимов, синонимов и т.д.), а также большая избыточность. Например, в русском языке величина избыточности достигает 72—88%. Если при речевом общении она облегчает понимание, то при поиске в ИС избыточность приводит к увеличению шума, уменьшению полноты и точности выдачи информации. Что же представляет собой информационно-поисковый язык?

Информационно-поисковым языком называется некоторая формализованная семантическая система, предназначенная для выражения смыслового содержания документов и информационных запросов с целью отыскания необходимых сведений.

В абстрактном виде ИПЯ состоит из списка символов; правил преобразования, т.е. правил составления комбинаций символов; правил интерпретации, т.е. как следует понимать эти выражения.

ИПЯ должен удовлетворять следующим требованиям: 1) быть однозначным, т.е. каждый символ должен иметь только одно толкование; 2) каждое понятие должно выражаться только одним словом или словосочетанием; 3) он должен допускать возможность

перевода с естественного языка на информационный и наоборот и формализацию процедуры информационного поиска.

Символами ИПЯ могут быть не только термины, взятые из естественного языка, но и систематизированные понятия в виде цифровых, алфавитных, математических и других подобных им символов.

Информационные языки можно классифицировать по разным признакам: наличию или отсутствию грамматических средств, методике составления словаря и т.д. Однако основным свойством, характеризующим тот или иной информационный язык, является его семантическая сила. Семантическая сила – это понятие сугубо относительное и применимо только к искусственному языку. Она показывает, насколько полно и точно искусственный язык способен выражать необходимую для поиска информацию.

В настоящее время для целей информационного поиска используют 2 основных типа ИПЯ: классификационные языки и языки дескрипторного типа. Вопрос о применении того или иного ИПЯ во многом зависит от конкретной информационно-поисковой системы.

Для описания документов в библиотеках, архивах обычно применяют классификационные языки. К ним относятся традиционные библиотечно-библиографические классификации, универсальная десятичная классификация, алфавитно-предметные указатели.

Языки этого типа имеют следующие особенности:

- эти системы являются двухступенчатыми. Первоначально по алфавитному указателю необходимо искать классификационный индекс, а затем, по индексу, и сам материал. Это существенно замедляет поиск;
- основной словарный состав задается в виде фиксированного списка слов, словосочетаний и фраз;
- структура списка строго фиксирована и его пополнение затруднено;
- язык имеет недостаточную семантическую силу для обеспечения информационного поиска по узкотематическим запросам потребителей.

Языки классификационного типа традиционно используются в ручных ИПС, особенно там, где информационные массивы отличаются большими масштабами и разнообразием. При создании классификационного языка используется логическая операция, связанная с классификацией информации по определенному принципу. Основная проблема здесь заключена в определении основания для классификации: это может быть предметно-отраслевой принцип, как,

например, в ББК (библиотечно-библиографическая классификация), либо видовой, как в управленческих документальных системах.

Классические ИПЯ – это системы классификации, широко применяемые в библиотечной практике. ББК возникли в древние времена. Например, в период средневековья господствовала классификационная схема, основу которой составляли «семь свободных искусств» – грамматика, диалектика, риторика, арифметика, геометрия, музыка и астрономия.

К концу XV в. появились так называемые факультетские классификации, отражавшие существующее в то время деление факультетов в университетах. В Европе университеты обычно имели философский, юридический, медицинский и теологический факультеты, в соответствии с этим классифицировалась и литература. Эти классификации просуществовали длительное время, вплоть до XIX в.

В России библиотечно-библиографические классификации начали разрабатываться в XIX в. Одну из схем составил П. Г. Демидов для каталога личной библиотеки. Он был ученым-натуралистом, поэтому в его схеме наиболее детализированы разделы естественных наук и технологии. В 1841 г. академик К. Э. Бэр разработал классификацию, в которой впервые сделал попытку расположить науки в последовательности, отражавшей историю развития мира: науки о неорганической природе, науки об органической природе и науки о человеке и человеческом обществе.

Как система ББК представляет собой логическое расположение индексированных рубрик по иерархической, или древовидной, схеме. В основе любой классификации лежит представление о том, что все накопленные человечеством знания могут быть разделены на взаимоисключающие классы и подклассы. Взаимное проникновение одних наук в другие и их взаимосвязь исключаются. Каждый класс делится лишь по одному признаку и дает серию подклассов. Количество классов и подклассов ограничено.

Существует ряд формально-логических правил, которым подчиняется любая ББК:

- 1) деление наук по классам и подклассам производится только по одному основанию;
- 2) подклассы должны исключать друг друга;
- 3) должна соблюдаться непрерывность при делении классов на подклассы и подклассов на понятия. Скачкообразные переходы не допускаются.

Основные виды отношений между элементами классификации – это иерархия и соподчинение. Под иерархией понимаются такие отношения, при которых один элемент системы является подклассом другого элемента, более широкого. Соподчиненность – отношения между подклассами одного и того же класса.

Термины в иерархических классификационных системах располагаются в порядке постепенного перехода от общих к более частным. ББК обычно издаются в виде таблиц, состоящих из 2 частей: основных таблиц и таблиц определителей. В основных таблицах все отрасли знания и их разделы расположены в логической последовательности. Каждая рубрика схемы содержит ряд подчиненных ей подрубрик и образует определенную иерархию.

Помимо предметных рубрик, предусматривающих деление документов по их содержанию, ББК включают ряд классификационных схем, которые дают возможность систематизировать материал по формальным признакам, например, виды изданий (книги, периодические издания, патенты и т.д.), по языку их текста, по стране издания. Каждому классу и подклассу, рубрике и подрубрике ББК присваивается условное обозначение, получившее название индекса. Они могут быть цифровыми, буквенными и смешанными. Преимущество цифровых индексов состоит в том, что порядковая последовательность цифр более очевидна, чем алфавитная последовательность букв. Недостаток цифровых индексов – ограниченность базы.

Индексы могут быть порядковые и ступенчатые. В первом случае при обозначении рубрик используются порядковые номера подразделений классификации в общем числовом ряду. В большинстве ББК применяются ступенчатые индексы. Они отражают ступенчатую структуру классификации, каждый ее раздел обозначается каким-либо одним символом, а все деления этого раздела – двумя символами, из которых первый обозначает раздел, а второй – соответствующий подраздел. Применение ступенчатых индексов позволяет проводить неограниченную детализацию классификационной схемы.

Наиболее известными в настоящее время иерархическими классификациями являются: десятичная классификация М.Дьюи, универсальная десятичная классификация (УДК), французская классификация и др. В России используется ББК, разработанная в библиотеке им. В.И.Ленина.

Наиболее распространенной ББК современности является УДК. В качестве базы этой схемы была принята классификация Дьюи. В

настоящее время УДК применяется более чем в 50 странах, с 1962 г. УДК была введена в нашей стране для всех органов НТИ и технических библиотек. УДК отличает:

- 1) охват всех отраслей знания;
- 2) использование только цифровой десятичной системы индексации, позволяющей производить неограниченную детализацию подклассов без нарушения основной структуры классификации;
- 3) применение только арабских цифровых индексов;
- 4) возможность индексирования документов с любой степенью дробности.

УДК состоит из 2 частей: таблиц основных индексов и вспомогательных таблиц общих и специальных определителей. Первая часть представляет собой таблицы, в которых вся совокупность человеческих знаний разделена на 10 основных разделов (классов), каждый из которых в свою очередь подразделяется на 10 подклассов, а подклассы имеют 10 более мелких подразделений. Основные отделы УДК следующие:

- 0 – общий отдел, библиография, библиотечное дело;
- 1 – философия, психология;
- 2 – религия, теология;
- 3 – общественные науки, право;
- 4 – отдел оставался свободным;
- 5 – математика, естественные науки;
- 6 – прикладные знания;
- 7 – искусство;
- 8 – художественная литература, филология, языкознание;
- 9 – география, история.

Например:

- 6 – прикладные науки;
- 62 – техника в целом;
- 624 – строительная техника;
- 624.1 – строительство, общие вопросы;
- 624.13 – земляные работы.

Точка ставится после каждой третьей цифры и способствует лучшей обозримости всего индекса. Фактически каждому понятию, указанному в основной таблице, соответствует свой индекс УДК. Например, 61.177 – перфорационные машины.

Кроме основных индексов в УДК используются общие и специальные определители, которые служат для более детальной характеристики индексируемой работы. Определители используются

для выражения признаков, общих для многих предметов, процессов, явлений. К общим определителям относятся: определители формы и характера материала ( 05 – периодические издания, 08 – энциклопедии), определители места (47 – СССР), определители времени (например, «1996» – 1996 г.), определители точки зрения (001 – теоретическая точка зрения, 002 – точка зрения реализации) и т.д. Специальные определители отражают признаки, характерные лишь для узкого круга предметов и явлений.

Кроме основных индексов, общих и специальных определителей в структуру УДК входят знаки соединений для более полного и точного отражения индексируемых материалов.

Например: «+» – присоединение (5+62 — естествознание и техника); «/» – распространение (546.32/34 — от калия до лития); «:» – отношение (016:621.3 – библиография по электротехнике); «'» – апостроф означает объединение (546.32'175 – нитрат калия).

Наряду с описанными библиотечно-библиографическими классификациями в библиотечной практике применяются так называемые фасетные классификации. В фасетных классификациях различные термины группируются по категориям, а порядок перечисления этих категорий определяется особыми приемами. Составители фасетной классификации рассматривают класс как однородный предмет, например, химия, физика, медицина, история и т.д. Подразделения внутри классов состоят из категорий, которые определяются на основе различных характеристик. В медицине, например, имеются такие категории, как орган, проблема (болезнь), симптом (лихорадка), агент (вирус), метод лечения (хирургия).

Одновременно устанавливается жесткий порядок, в котором должны располагаться категории. Например, на первом месте будет стоять орган, на втором – проблема, на третьем – симптом и т.д.

Процедура индексирования при фасетной классификации сводится к следующему: изучается документ и на естественном языке формулируется его краткое содержание. Затем это содержание переводится в поисковый образ документа с использованием фасетных индексов.

Например, статья, в которой сообщается о запуске реактора на тепловых нейтронах, с помощью фасетной классификации выглядит следующим образом: R212.2D<sub>2</sub>O-081.2 , где R2 – реакторы, R212.2 - реакторы на тепловых нейтронах, D<sub>2</sub>O – тяжелая вода, 081.2 – обогащенный уран.



Фасетные классификации имеют ряд преимуществ перед иерархическими системами: они позволяют производить многоаспектное индексирование документов, допускают большую глубину деления понятий, чем традиционные классификации, в эти системы значительно легче включить новые термины. Однако фасетные классификации имеют и ряд недостатков – это трудоемкость построения такой классификации, поэтому они создаются только для узкоспециализированных информационных систем.

Создание новых или переработка старых классификаций — дело трудное. Поэтому еще в 70-е гг. активно стали разрабатываться методы автоматизированной разработки классификационных схем. Автоматические методы составления классификаций целесообразно применять для таких документов, где обычные ББК неприменимы (архивная документация, управленческая документация и т.д.).

Автоматическое составление классификации предполагает несколько этапов:

- анализ массива документов методом математической статистики с целью определения числа рубрик и формулировки их названий для создаваемой классификации;
- выявление в анализируемом тексте ключевых слов, сравнение их с терминами классификации;
- составление поискового образа документа в терминах классификационной схемы.

Второй тип информационно-поисковых языков (дескрипторный) чаще используется в автоматизированных и механизированных ИПС. Он основан на методе координатного индексирования и состоит из дескрипторов и ключевых слов. Дескриптор – это любая лексическая единица языка, выражающая общее понятие и используемая для описания основного смыслового содержания документов. Дескрипторы могут быть представлены в виде слов естественного языка (ключевые слова), индексов, кодов. Под ключевыми словами понимаются такие слова естественного языка, которые взяты из текста документа и наиболее полно раскрывают его содержание. Индекс – код – это образ знака или слова естественного языка, выраженные в символах какого-либо искусственного языка.

В случае координатного метода индексирования поисковый образ документа состоит из набора ключевых слов, расположенных в алфавитном порядке. Чтобы составить поисковый образ документа, достаточно выбрать из его текста те слова, которые наиболее полно характеризуют документ, и расположить их в алфавитном порядке.

Таким же путем составляется и поисковое предписание. Достоинством координатного метода индексирования является то, что этот метод позволяет использовать в поисковом образе документа неограниченное число терминов, что в свою очередь ведет к сокращению потери информации. Каждый термин выбирается свободно из текстов документов.

Дескрипторные ИПЯ могут быть организованы как по иерархической системе в виде тезаурусов, так и в виде обычных алфавитных словарей.

Из отобранных ключевых слов составляется словарь информационного языка. Эти слова подвергаются тщательной лексикографической обработке – устраняются синонимы, омонимы. Таким образом мы получим дескрипторный словарь. В словаре дескрипторы пишутся в именительном падеже и, как правило, в единственном числе, ключевые слова, принадлежащие к другим частям речи, – прилагательным, глаголам, должны быть преобразованы в существительные. Обычно в дескрипторном словаре дескрипторам присваиваются цифровые индексы, словосочетания могут обозначаться как простыми, так и составными индексами. Узкотематические дескрипторные словари без грамматики обеспечивают высококачественный поиск информации особенно в специализированных системах. Если же ИПС рассчитана на справочно-информационные фонды с большим объемом информации и широкой тематикой, то точность поиска можно обеспечить только за счет увеличения количества дескрипторов в поисковых образах документов, но это ведет к повышению «поискового шума».

С целью обеспечения оптимального координатного индексирования создаются специальные словари-справочники, в которых дескрипторы объединяются в смысловые группы. Такие словари называют тезаурусами. Слово «тезаурус» в переводе с греческого означает «хранилище», «запас», «сокровищница». Тезаурус – это такой словарь-справочник, в котором перечислены дескрипторы и синонимичные им ключевые слова, показаны методы устранения синонимии, омонимии, полисемии, выражены родо-видовые и ассоциативные связи дескрипторов. Составление тезауруса отнимает много времени, а последующие изменения затруднены.

Тезаурус состоит из 3 основных частей: классификационной схемы, алфавитного перечня дескрипторов и ключевых слов и тематических полей. Классификационная схема позволяет быстро найти ту группу дескрипторов, которая связана с интересующими индексатора

вопросами. Число уровней в классификационной схеме определяется количеством терминов, требованиями объема и областью применения системы. Алфавитный перечень дескрипторов и ключевых слов – это самая большая по объему часть тезауруса. Она представляет собой словарь дескрипторов и их синонимов. Тематические поля позволяют находить максимальное количество дескрипторов для определенной области науки. Тематические поля соответствуют разделам классификационной схемы, но связи выходят за рамки разделов. Пользование такими схемами дает возможность быстрее индексировать документы.

Таким образом, дескрипторный язык – это ИПЯ, у которого словарь состоит из перечня дескрипторов, а грамматика представляет собой набор правил, используемых для образования поисковых образов документов и поисковых предписаний.

Дескрипторные языки более удобны для узкопредметного тематического поиска; обеспечивают большую точность поиска; позволяют дополнять лексический состав языка новыми терминами.

Исследования в области ИПЯ в последние годы показали, что дальнейшее развитие информационных языков должно пойти по линии соединения преимуществ дескрипторных языков с положительными сторонами классификаций.

### **3.6. Показатели эффективности информационной системы**

Целью системы является удовлетворение информационных потребностей потребителей. Поэтому эффективность ИС практически всегда оценивается через критерий полноты удовлетворения этих потребностей.

Выделяется техническая и экономическая эффективность ИС. Техническая эффективность – это мера способности системы выполнять те функции, для которых она предназначена, а экономическая эффективность служит мерой стоимости выполнения этих функций.

Основные факторы, влияющие на величину затрат при работе системы:

- первоначальные затраты на проектирование и внедрение ИС;
- затраты на приобретение необходимой техники, помещений, подготовку кадров;
- объем первоначального массива информации;
- скорость приращения исходного массива за единицу времени;
- логическая сложность задач, решаемых с помощью данной ИС;
- количество запросов в единицу времени;

- способы выдачи и передачи информации на расстояние;
- трудоемкость подготовки материалов для занесения в ИС;
- форма представления конечных результатов.

Для определения технической эффективности используется ряд параметров, связанных понятием релевантности. Под релевантностью понимается отношение количества характеристик поискового образа записи или документа к общему числу характеристик, составляющих поисковый запрос. Документ, основное смысловое содержание которого отвечает на информационный запрос, называется релевантным. Оценка релевантности является фундаментальным процессом информационного поиска. Соответственно выделяются такие параметры ИС, как:

1) коэффициент полноты (отношение числа выданных релевантных документов или записей к их общему числу в поисковом массиве)

$$p = \frac{r}{N} * 100\%,$$
 где  $p$  – коэффициент точности,  $r$  – релевантная информация в выдаче;  $N$  – общее число релевантных документов и записей в массиве;

2) коэффициент точности (отношение релевантных документов и записей в выдаче ( $r$ ) к общему числу их, полученному в результате поиска ( $S$ )).

$$T = \frac{r}{S} * 100\%.$$

Между полнотой и точностью информационного поиска существует обратная зависимость. Рабочими характеристиками системы считаются: коэффициент полноты – от 80 до 100%, коэффициент точности 60—90% ( в 1970-е гг. данный показатель колебался от 8 до 20%);

3) коэффициент шума – отношение нерелевантных документов в выдаче к общему числу документов, полученных в результате поиска.

$$E = \frac{S - r}{S} * 100\%,$$
 где  $S$  – всего документов в выдаче;  $r$  – количество релевантных документов в выдаче.

#### **Глава 4. Интегральные и экспертные информационные системы**

Всесторонняя автоматизация научно-информационной деятельности на базе использования последних достижений науки и техники, а также усложнение потребностей информационного

обеспечения привели к созданию и внедрению в информационную практику автоматизированных информационных систем особого типа, получивших название интегральных информационных систем (ИИС). ИИС – это совокупность методов и средств, позволяющих при однократном описании, индексировании и реферировании документов и преобразовании их в машиночитаемую форму обеспечивать многоаспектную их обработку и многофункциональное использование для удовлетворения самых разнообразных информационных потребностей.

Интегральные системы в первую очередь стали разрабатываться для нужд научно-информационной деятельности. В частности, с 1971 г. в ВИНТИ начал реализовываться проект, целью которого было создание Государственной автоматизированной системы научно-технической информации (ГАСНТИ). Позднее аналогичная система была создана в ИНИОН.

В сущности, интегральная система – это комплекс, состоящий из нескольких более простых подсистем, каждая из которых предназначена для выполнения какой-то одной функции. Взятая в единстве, такая система обладает способностью удовлетворять любые типы информационных потребностей, в то время как простые ИПС ориентированы на удовлетворение информационных потребностей только определенного типа.

Так, например, в интегральной системе научно-информационного профиля должны быть реализованы следующие функциональные подсистемы:

- избирательное распространение информации;
- подготовка бюллетеней сигнальной информации;
- подготовка реферативных журналов и указателей;
- ретроспективный поиск документов и информации.

Для создания подобных систем технически должна быть обеспечена возможность накапливать большие объемы информации и обеспечивать быстрый ее поиск. Такие условия возникли лишь после появления ЭВМ с запоминающим устройством большой емкости, соответствующими вводными и выводными устройствами.

Для исчерпывающей аналитико-синтетической обработки каждого документа в ИИС применяются специальные формуляры, которые заполняются при составлении библиографического описания документа, его индексирования, аннотирования и реферирования. Формуляр содержит следующие сведения о документе:

- рабочий номер;

- сведения о типе (статья, монография, патентное описание и т.д.);
- заглавие в переводе на русский язык;
- полное библиографическое описание;
- классификационные индексы;
- индексы УДК;
- список ключевых слов.

Этих сведений достаточно для подготовки информационных изданий любого типа и удовлетворения разнообразных информационных запросов. Благодаря накоплению в памяти компьютера текстов рефератов, аннотаций и библиографических описаний вместе с их индексами резко сокращается время подготовки информационных бюллетеней и реферативных журналов. ЭВМ по заданной программе выбирает из массива соответствующие тексты, располагает их в должном порядке и воспроизводит на выводном устройстве.

Итак, в интегральной системе нашел свое воплощение принципиально новый подход к созданию автоматизированных поисковых систем, который позволяет обеспечить максимальную экономию интеллектуального труда за счет исключения повторных операций.

В настоящее время интегральные системы созданы во всех информационных центрах мира. Подробнее познакомимся с аналогичной системой, созданной на базе Института научной информации по общественным наукам во взаимодействии с другими центрами информации и научными институтами. ИНИОН осуществляет самое полное в мире библиографическое издание в области общественных наук, охватывающее более 250 тыс. книг и статей, информация о них помещается в 30 ежемесячных указателях новой отечественной и зарубежной литературы, а также в многочисленных ретроспективных библиографических указателях. Кроме того, публикуются два реферативных журнала, включающие 16 серий, десятки реферативных сборников, экспресс-информация и сотни аналитических обзоров. Эти издания широко используются в нашей стране и распространяются более чем в 50 государствах мира. Такой масштаб работы невозможен без автоматизации всех видов научно-информационной деятельности, без создания интегральной информационной системы.

В 1980—1982 гг. были созданы первые отраслевые базы данных по экономике, демографии, философии, социологии. В 1985 г. к ним

добавились базы данных по истории, археологии, этнографии, науковедению, литературоведению, государству и праву. Помимо отраслевых начали функционировать и проблемные БД. С 1985 г. был осуществлен переход на автоматизированную подготовку библиографических указателей, что позволило вместо 70—75 тыс. документов в год обрабатывать 200—220 тыс.

В базах данных ИНИОНа в начале 1990 г. насчитывалось более 1,1 млн документов по проблемам экономики, философии, социологии, демографии, истории и по другим отраслям знаний, а также по комплексным проблемам социально-экономического и политического развития стран Азии, Африки, Латинской Америки.

Ежегодно база данных пополняется почти 220 тыс. новых документов. В ВИНИТИ объемы обрабатываемой информации еще более впечатляющие, здесь ежегодно вводится в научный оборот 1,5 млн документов. Базы данных ИНИОНа и ВИНИТИ доступны в режиме «он-лайн» как отечественным, так и зарубежным потребителям.

Создание и эксплуатация больших баз данных требует значительных затрат труда, времени и денег. По некоторым оценкам в настоящее время в базах данных в мире содержится свыше 300 млн документов, введенных за последние 10—15 лет. Средняя стоимость обработки, ввода и хранения одного документа составляет 30 долларов, т.е. общие затраты составили около 9 млрд долларов<sup>1</sup>.

Эффективность подобной информационной системы зависит от доступности ее информационных богатств потребителю. Поэтому важным моментом являются создание и реализация концепции сети органов информации, которая на различных уровнях и на различных организационных принципах взаимодействует с зарубежными и международными системами.

Сеть автоматизированных центров научной информации по общественным наукам (САЦНИОН) входит в государственную автоматизированную систему научно-технической информации. САЦНИОН – это сеть со структурой типа «звезда», в центре которой находится автоматизированный информационный центр ИНИОН. В состав САЦНИОН входят секции Президиума Российской Академии наук, вузы, министерства и ведомства ( всего по Москве 37 абонентов), региональные центры научной информации в Петербурге, Новосибирске, Ростове-на-Дону, Саратове, Екатеринбурге. На

---

<sup>1</sup> Подробнее см.: Региональные автоматизированные информационные системы по общественным наукам: Опыт проектирования и эксплуатации. Л., 1990.

международном уровне САЦНИОН дополняется автоматизированными центрами европейских стран, а также Кубы, Вьетнама, Монголии.

С помощью сети осуществляется передача данных в системе, облегчаются процессы коммуникации, улучшается организация обработки и хранения данных.

Организационно информационная сеть состоит из территориально рассредоточенных, взаимосвязанных информационных центров и подразделений. В ее структуре выделяются 4 уровня: международный, национальный, региональный, локальный.

Организации—участники сети составляют три группы по признаку приоритетности доступа к базам данных автоматизированного информационного центра ИНИОН. В первую группу входят институты и секции Президиума Академии наук. Они не только пользуются наивысшим приоритетом доступа, но и участвуют в формировании собственных локальных автоматизированных баз данных. Это чаще всего ретроспективные базы данных, которые создаются по согласованной тематике и единым методологическим принципам. Вторая группа представлена региональными информационными центрами, также участвующими в создании распределенного банка данных и формирующими свои локальные сети. Третья группа объединяет пассивных потребителей информации.

В рамках реализации программы создания распределенного банка данных были созданы локальные базы данных: «Все о Латвии» в ЦНИОН АН Латвии; по археологическим объектам в ЦНИОН АН Грузии; «Крупный город» в Институте социально-экономических проблем РАН в Санкт-Петербурге; «Восток» в Институте Востоковедения РАН и Институте Латинской Америки РАН. Использование информационных ресурсов локальных баз данных сталкивается с проблемой организации удаленного доступа сетевых пользователей к микрокомпьютерным базам данных.

В процессе развития сети возникают разнообразные технологические и организационные проблемы, в том числе проблемы децентрализованной обработки и ввода информации, качества отбора и обработки литературы, лингвистические проблемы, связанные не только с формированием БД, но и доступом к базам разноязычных пользователей, а также вопросы совместимости систем в части математического обеспечения и технических средств.

Дальнейшее развитие сети связано с совершенствованием системы управления, созданием диспетчерской службы, повышением



надежности сети передачи данных, организацией подготовки кадров и обучения потребителей.

Среди наиболее значимых задач, решаемых с помощью сети, назовем:

- кооперацию и разделение труда между информационными центрами;
- расширение круга абонентов;
- целенаправленное избирательное оповещение ученых о новых поступлениях;
- поиск документов в режиме ИРИ и ретроспективного поиска;
- создание интегрированного банка данных библиографических описаний и рефератов наиболее важных опубликованных в мире работ и фактографической информации;
- повышение качества и оперативности выпуска информационных изданий.

Создание экспертных систем (ЭС) представляет еще один перспективный путь развития автоматизированных информационных технологий. Под экспертной системой понимается такая ИС, которая объединяет возможности компьютера со знаниями и опытом эксперта в такой форме, что она может предложить разумный совет или осуществить решение поставленной задачи. Фактически такие ИС предназначены для получения ответов на заранее не предусмотренные вопросы, а также для получения данных и фактов, которые непосредственно не вводились в используемый информационный массив, путем их логического вывода из сведений, имеющихся в данном массиве. По сложности устройства и трудности построения решаемых задач такие системы весьма значительно отличаются от обычных фактографических автоматизированных систем и относятся к более высокому классу.

Первым шагом к созданию экспертных систем явилась разработка информационных языков и логики отдельных наук, что связано с анализом системы понятий, методов исследования, принятых в данной отрасли знания. Чем более формализован язык науки, тем проще создать для нее информационный язык, например, для математики, химии, медицины и т.д.

Одним из первых людей, пытавшихся создать своеобразную логическую машину – прообраз современных экспертных систем, был испанский философ и богослов Раймунд Луллий (1235 – 1315). Он выдвинул идею, что путем механического комбинирования понятий можно получить новую информацию. Он был убежден, что в каждой

отрасли знания имеется небольшое число простых основополагающих понятий и умозаключений – категорий, которые должны приниматься без доказательства. Если из этих категорий построить все возможные сочетания, то знания данной отрасли будут исчерпаны. Реализуя эту идею, он создал логическую машину, состоящую из нескольких кругов, вращаемых вокруг одного центра. На каждом круге были написаны слова, выражающие понятия. Вращая по определенным правилам один круг относительно другого, можно было добиться построения умозаключений, отражающих истину. Эти идеи получили достаточно широкое распространение в научной среде, в частности оказали большое влияние на Лейбница, Декарта и др. И сейчас практически никто не сомневается в способности ЭВМ решать достаточно сложные логические задачи.

В настоящее время экспертные системы получили широкое распространение в таких отраслях, как медицина, геология, химия и т.д. Эти системы могут диагностировать болезни, определять места добычи полезных ископаемых и решать другие задачи.

По поводу экспертных систем высказывается довольно много различных суждений и, как не странно, общая ошибка заключается в недооценке их возможностей. Специфика экспертных систем состоит в том, что они основаны на правилах логического вывода. В отличие от обычных компьютерных программ, реализующих принцип последовательного выполнения необходимых команд, экспертные системы включают набор правил, которые не предусматривают последовательность выполнения, а срабатывают только так и тогда, когда выполняются соответствующие условия.

Структурно информационный массив ЭС состоит из трех основных блоков: базы данных; базы знаний, представляющей модель знания эксперта; и блока логического вывода. База знаний содержит постоянную информацию, неизменные данные, которые характеризуют предметную область системы в целом. Чаще всего знания эксперта представлены в виде особых правил, имеющих логическую структуру «если... то...». База данных выполняет другую задачу, она содержит переменные сведения, ее специфика связана с задачей, выполняемой на данный момент.

В 1965 г. была создана одна из первых экспертных систем, которая получила название DENDRAL. Она появилась в Стэнфордском университете и представляла собой результат совместных усилий специалистов по компьютерам и группы экспертов в области химии. Ее основная задача – автоматизировать процесс определения правильной

химической структуры вещества. Пользователь дает системе некоторую информацию о веществе, а также данные спектрального анализа, на основе чего ЭС строит соответствующую химическую структуру. Описание структуры решается с помощью графических средств, а также с помощью понятийного языка науки. Основная сложность здесь состоит в том, что число возможных структур вещества может быть достаточно большим. Система решает эту проблему путем введения ограничений – выдает за один цикл возможные химические структуры, а затем их последовательно проверяет, сверяя свои гипотетически полученные результаты спектрального анализа с настоящими результатами, чтобы определить наиболее верную структуру. По результатам работы данной системы было написано более двух десятков научных работ.

Другим примером информационной системы такого рода является медицинская экспертная система MYCIN, предназначенная для диагностики и лечения заражения крови и менингитных инфекций. Система ставит соответствующий диагноз, исходя из представленных ей симптомов, и рекомендует курс медикаментозного лечения. База знаний данной системы состоит в общей сложности из 450 правил, разработанных с помощью специалистов по инфекционным заболеваниям Стэнфордского университета.

Основополагающим принципом, на котором строится решение предлагаемых задач в рассматриваемой экспертной системе, является принцип использования вероятностного подхода. И это закономерно, так как медицинская диагностика – наука неточная, она во многом основана на интуиции, так как один и тот же набор симптомов может иметь различное происхождение. Система MYCIN справляется с диагностикой путем выявления всех возможных диагнозов и расчета показателя определенности. В результате система выдает целый список диагнозов, называя показатель определенности для каждого из них, и врач, основываясь на своем опыте, выбирает тот диагноз, который кажется ему наиболее обоснованным. Математически это выглядит несколько сомнительно, но на практике срабатывает хорошо.

Гуманитарные и общественные науки в меньшей степени поддаются формализации, поэтому разработка информационного языка и экспертных систем в данных областях началась значительно позднее и не получила еще достаточно широкого распространения. Интересную экспертную систему для проведения топонимических исследований

разработал Ю.Е.Храмов в 1991 г.<sup>2</sup> Она получила название ГИДРОНИМИКОН и предназначена для анализа гидронимов Восточно-Европейской равнины. В качестве программной оболочки было использовано программное средство ЛОТА, имеющее блок распознавания образов и процедуры логического вывода.

Гидронимы представляют собой наиболее древнюю и устойчивую часть топонимов – названий географических объектов и включают названия рек, озер, болот. Анализ каждого гидронима – трудоемкая задача, требующая высокой квалификации, привлечения огромной суммы знаний лингвистов, историков, археологов. Однако их изучение позволяет раскрыть целый спектр научных проблем, связанных с передвижением древних народов, прародиной славян и т.д.

Для построения базы знаний были использованы работы ведущих специалистов в области топонимии и исторического языкознания, использовано более 800 монографий и статей. ГИДРОНИМИКОН включает обширный справочный материал по археологическим культурам, истории языка и народов большинства рассматриваемых регионов.

Анализ гидронимов происходит путем поиска фонетически сходных названий. В этих целях система располагает весьма обширной базой данных по гидронимам Восточной Славии, Прибалтики, Западной Сибири, Польши, Балкан, Средней и Западной Европы, включая и исторические варианты названий. При поиске также используется разработанный Ю.Е.Храмовым алгоритм сопоставления образов. Средства экспертной системы позволяют картографировать найденные примеры.

Основные функции ГИДРОНИМИКОНА – это анализ гидронимов с подготовкой текстового материала, обосновывающего выводы; подбор и картографирование примеров лингвистических явлений; выдача подробных справок по источникам и библиографии.

Развитие технологии экспертных систем с использованием правил логического вывода, концепции теории нечетких множеств и других создают более полное и адекватное представление знаний эксперта даже о тех областях науки, которые в меньшей степени подвержены формализации. Такой подход, в частности, реализован в системе АМСОР, созданной в 1986 г. в МГУ (автор Л.И. Бородкин).

---

<sup>2</sup> См.: Храмов Ю.А. ГИДРОНИМИКОН – экспертная система по гидронимии Восточно-Европейской равнины // Информ. бюл. Комиссии по применению математических методов и ЭВМ в исторических исследованиях. 1992. №5.

Возможности данной системы были использованы для решения задач аграрной типологии Европейской России в конце XIX – начале XX в.

## **Глава 5. Информационные технологии в архивном деле**

В современной публицистике архивы часто называют сокровищницей недоступных знаний о важнейших событиях прошлого, подчеркивая тем самым факт, что информационные технологии большинства архивов в нашей стране до сих пор основываются на приемах и методах обработки документов, сложившихся еще в середине XIX в.

Архивным работникам приходится иметь дело с невероятно большим потоком информации, осуществляя операции по организации хранения документов, обеспечению их сохранности, учету, описанию, поиску. Ими просматриваются и описываются миллионы дел. Объем государственного архивного фонда достиг более чем 350 млн единиц хранения и постоянно растет. В этих условиях особое внимание уделяется поддержке в актуальном состоянии справочников и учетных документов.

Для осуществления операций с документами в архивах сложилась система научно-справочного аппарата, который состоит из совокупности учетных документов и архивных справочников. Учетные документы выполняют функцию закрепления, фиксирования принятой в отечественных архивах классификации источников, при решении вопросов фондирования, комплектования и обеспечения сохранности архивных фондов. Архивные справочники предназначены для поиска необходимых сведений об архивах, фондах, делах, документах. Любой архивный справочник можно назвать информационно-поисковой системой документального типа. Основой архивной ИС является совокупность описаний, т.е. той информации, которая в сжатой форме характеризует состав и содержание архива, фонда, дела. Для удобства поиска вся совокупность описаний в архивных справочниках классифицируется на группы по хронологическому, предметному и другим признакам, образуя классификационную схему.

Исходя из объекта описания справочника, выделим четыре основных вида архивных ИС:

- справочник по архивам (уровень описания – архив);
- путеводитель архива (уровень описания – фонд);
- опись дел (уровень описания – дело);

– каталог документов (уровень описания – документ).

Основная часть архивных справочников создана в виде книг и картотек и относится к ручным традиционным информационным системам. Наряду с ними со второй половины 70-х гг. в нашей стране были начаты работы по созданию автоматизированных информационных систем.

Характеризуя основные направления развития информационных технологий в архивном деле, выделим два основных направления: 1) совершенствование традиционного научно-справочного аппарата; 2) внедрение автоматизированных систем в научно-информационную деятельность архивов.

Справочники по архивам содержат информацию о существующих архивах и архивных службах в регионе, стране, мире. Они издаются, как правило, органами информации. Наиболее полным источником сведений об архивах в мире является «Международный справочник об архивах», опубликованный в 22 – 23-м томах «Архивума» в 1975 г. Данный справочник содержит информацию о более чем 2500 архивах 132 стран мира. В него включены библиографические данные и справки об основных фондах архивов. Наряду с данным справочником издан также ряд региональных изданий по странам Африки, Азии и др.

Более полными по составу информации являются национальные справочники и путеводители по архивам. Они различны по широте представленных в них сведений. Иногда они включают данные только о государственных архивах, но часто встречается информация об архивохранилищах других категорий, в том числе муниципальных, музеев, библиотек, церквей, а также о семейных архивах (напр., справочники «Государственные архивы Румынии», «Путеводитель по архивам Венгрии» и др.). Национальные путеводители, как правило, красочно оформлены и используются в туристическом бизнесе. Справочники дают первоначальное общее представление о распределении фондов архивов, их содержании.

Основными справочными средствами для исследователя были и остаются путеводители архива и описи дел. Они являются традиционными и издаются с XIX в. Архивы постоянно работают с описями и путеводителями, совершенствуют и обновляют их. Большое внимание переизданию путеводителей и описей уделялось в 50-е гг. в связи с необходимостью уточнения состава документов в результате потерь в период Великой Отечественной войны. Однако если обновление и издание путеводителей по архиву осуществляется с определенной периодичностью, то публикация описей дел остается

достаточно редким событием. В большинстве отечественных архивов существует несколько экземпляров описей, подготовленных в рукописном варианте, и познакомиться с ними можно только в читальном зале архива.

Справочники этого вида готовятся силами конкретного архива, поэтому наблюдаются существенные различия в степени детализации информации. Это может быть или просто список фондов, или подробное их описание. Полноценный справочник обычно включает развернутое описание фондов, объемом 1-2 стр. В описании фонда должны быть представлены: история фондообразователя; общее количество документов; крайние их даты; список утраченных материалов с указанием документов, их восполняющих. Для крупных архивов наряду с полными путеводителями часто издается их сокращенный вариант.

Большой интерес для исследователей представляет издание тематических справочников документов, посвященных наиболее важным историческим событиям (Октябрьская революция, Великая Отечественная война и др.), которые охватывают материалы нескольких хранилищ.

Наряду с указателями и справочниками большое внимание уделяется созданию каталогов. Систематические каталоги имеют практически все государственные архивы. Они представляют собой важнейший инструмент интенсивного использования документов. Работа по созданию систематических каталогов началась в 60-е гг., однако в силу ряда причин качественный уровень был невысоким. В 70-е гг. каталогизация становится плановой и осуществляется вместе с внедрением «Схемы единой классификации документной информации в систематических каталогах государственных архивов СССР». Многие архивы в этот период стали проводить экспертизу ценности каталогов, которая выявила существенные недостатки: неразработанность многих важных разделов, несоответствие каталожных шифров состоянию единиц хранения и т. д. В результате было уничтожено до 30% объемов каталогов.

Архивные каталоги, являясь межфондовыми справочниками, позволяют вести поиск одновременно по всем включенным в каталог фондам. Ядром системы каталогов для большинства крупных и средних архивов является системный каталог, который в зависимости от специфики архива может дополняться именовым, административно-территориальным, каталогом по истории государственных учреждений и др. Создание полноценной системы каталогов – сложная и очень трудоемкая задача. Свидетельством этого является тот факт, что в

настоящее время, по оценкам специалистов, каталоги отражают информацию фондов лишь на 15%. В первую очередь они используются сотрудниками архивов для выполнения работ по запросам, а также исследователями. Каталогизация архивов выступает как одно из основных направлений совершенствования научно-справочного аппарата, вероятно, поэтому внедрение автоматизированных технологий охватило в первую очередь именно эту сферу.

В СССР работы по созданию архивных автоматизированных информационных систем начались в середине 70-х гг. и велись в двух направлениях: 1) разработка автоматизированного каталога на базе Центрального фондового каталога Главархива Российской Федерации (АСНТИ ГАРФ); 2) создание АИПС на отдельные тематические комплексы архивных документов.

Автоматизированная система научно-технической информации по материалам ГАРФ была введена в эксплуатацию в 1980 г. Она предназначалась для поиска информации о фондах, содержащих документы по заданной теме; подготовки и издания межархивных справочников; автоматизации процессов государственного централизованного учета архивных фондов в масштабах страны. Информационный массив системы формировался из описаний фондов, карточки описания составлялись на местах работниками архивов. В первую очередь были введены сведения о фондах органов государственной власти. На каждый архивный фонд составлялась аннотация о составе документов и их содержании. Аннотация подвергалась анализу, в результате которого выделялись ключевые слова, в совокупности составлявшие поисковый образ фонда. Каждое поле карточки имело строго определенное место и заполнялось с соблюдением правил. Карточки поступали в Главархив РФ, где они регистрировались и вводились в систему.

На основе поисковых образов карточек после соответствующей лингвистической обработки был составлен словарь ключевых слов, преобразованный в информационно-поисковый тезаурус системы. Для тематического поиска применялся также рубрикатор АСНТИ, его рубрики соответствовали основным тематическим направлениям использования документов архивных фондов.

Планировалось, что АСНТИ станет базовой автоматизированной системой государственной архивной службы, однако в полной мере система так и не была реализована. В 1984 г. были проведены региональные совещания по АСНТИ и принято решение разработку системы отложить. Причинами такого решения можно назвать



недостатки технической оснащенности системы, информационно-поискового языка, отсутствие эффективных коммуникаций и др.

Наряду с этим проектом в 80-е гг. предпринимались попытки создания автоматизированных поисковых систем по тематическим комплексам документов, в частности по истории Октября. Основная задача подобных АИПС состояла в обслуживании потребителей в режиме ретроспективного поиска и избирательного распространения информации, а также подготовке справочных информационных изданий.

Поисковая система на комплекс «Документы Октября» включала около 130 тыс. описаний архивных документов, которые хранятся в более чем 100 архивах страны. Однако основная часть этих документов была ранее уже опубликована и введена в научный оборот, поэтому эффективность системы оказалась низкой, ее деятельность ограничилась изданием тематического каталога и топографического указателя, в котором информация группировалась по архивам, хранящим документы.

Более удачным оказался другой опыт, связанный с созданием аналогичной информационной системы по комплексу документов «Памятники архитектуры и градостроительства». Система включала примерно 110 тыс. описаний документов, относящихся к досоветскому периоду (начиная с XVII в.), в том числе графические материалы. Общее количество запросов от архитектурных, проектных и строительных организаций на эти документы было очень большим, так как они необходимы для составления планов застройки, проектов реставрации и др.

Для лингвистического обеспечения системы использовался ИПЯ дескрипторного типа. Специфика тезауруса «Архитектура и градостроительство» состоит в том, что основная его часть представлена именами собственными — топонимами, именами архитекторов, строителей, владельцев домов. При составлении словаря необходимо было учесть исторический контекст существования объектов. Архитектурный памятник мог принадлежать в разные периоды разным владельцами, мог менять адрес и т.д. Так, например, дворец Лефорта был передан во владение А.Меньшикову, некоторое время принадлежал царской семье, затем — дворцовому ведомству. Естественно, что вся история существования памятников не может быть выражена с помощью простых средств. Каждый случай рассматривался индивидуально, чаще всего предпочтение отдавалось наиболее известному наименованию. Последняя редакция словаря содержит

около 8 тыс. дескрипторов и их количество продолжает расти. Информационно-поисковый тезаурус имеет рубрикатор, он построен по географическому принципу и содержит 666 рубрик, из которых 636 относятся к Москве, остальные – к Петербургу.

В качестве основных ориентиров при поиске документа служат: название архитектурного объекта, его местонахождение, фамилия архитектора. Выявление документов для АИПС «Архитектура и градостроительство» проводилось последовательно по фондам и коллекциям архивов и музеев с обязательным просмотром дел. На каждый документ заполнялся рабочий лист, информация которого затем переносилась в ЭВМ. Рабочий лист содержал следующие сведения:

- архив, номер фонда, описи, дела;
- местонахождение и тип объекта, его описание;
- вид документа;
- подлинность;
- способ воспроизведения (чернила, карандаш);
- заголовок дела и его аннотация.

Нужно отметить, что в результате деятельности системы не только осуществлялся поиск по запросам, публиковались географические, предметные и именные указатели, она позволила получить новую информацию: идентифицировать архитектурные памятники, установить принадлежность дома разным владельцам, дополнить биографии архитекторов и т.д.

Внедрение персональных компьютеров в деятельность архивов, появление прикладных программ, обеспечивающих диалог с машиной, явились своеобразными толчками к созданию автоматизированных каталогов на местах. Оператор вводит в машину подготовленную архивистом карточку с описанием документа. Подобное описание может включать следующие сведения: вид документа; содержание документа (аннотация); автор; корреспондент; время и место составления; опубликование; внешнее оформление; местонахождение (архивный шифр).

Использование схем классификации документной информации, а также ключевых слов позволяет обеспечить быстрый и точный поиск информации. Такой комбинированный поиск означает, что система в отношении массива документов, исполняет функции 5 традиционных каталогов – систематического, предметного, именного, географического и хронологического, т.е. происходит многократное увеличение поисковых возможностей научно-справочного аппарата.

Этап информатизации архивов переживают такие страны, как США, Великобритания, Франция, Германия, Канада и др. Наиболее высокоорганизованная сеть архивных АИС существует в США. В нее входят не только собственно архивные учреждения, но и библиотеки, архивохранилища ряда университетов.

Архивные автоматизированные сети имеются в Германии, они включают \* общинные, церковные, поместные архивы. В Великобритании, Канаде созданы локальные информационные сети, которые обеспечивают контроль за поступлением новых материалов, поиск информации, создание учетных и справочных средств в госархивах.

В нашей стране в последние годы происходит во многом стихийное развитие информатизации государственных архивных учреждений, что связано с отсутствием единых программ и принципов автоматизации деятельности архивов, существенными различиями в обеспеченности компьютерной техникой. Автоматизация архивного дела имеет главной своей задачей развитие системы формирования, обеспечения сохранности, всестороннего использования архивного фонда и широкого доступа к информационным ресурсам.

Почти все архивы, располагающие компьютерами, начали создавать базы данных, как тематические, обеспечивающие использование документов, так и по другим направлениям деятельности архивов. По оценкам специалистов более 88% существующих в настоящее время баз данных в архивах занимают тематические ИС; 10% – БД учетного характера; около 2% -административно-управленческие. Наиболее активно в процесс автоматизации включились госархивы федерального уровня.

Внедрение и активное использование автоматизированных архивных технологий должны сделать труд архивиста более продуктивным и содержательным, избавив его от выполнения многих рутинных операций.

Одним из актуальных вопросов деятельности архивов является работа с электронными документами. К настоящему времени в учреждениях и в организациях накоплен значительный массив документов на магнитных носителях, которые также должны приниматься на хранение. Основная сложность здесь состоит в том, что машиночитаемые документы в отличие от документов на традиционных носителях неразрывно связаны с программным обеспечением, с помощью которого они создавались. Возникает необходимость разработки и осуществления специальной программы,

предусматривающей организацию специализированных архивных учреждений или подразделений, укомплектованных квалифицированным персоналом, с целью создания нормативно-методической базы по работе с машиночитаемыми документами. Большую работу в этом направлении проводит Всероссийский научно-исследовательский институт документооборота и архивного дела (ВНИИДАД), представивший концепцию автоматизированной архивной технологии (ААТ).

Основные направления автоматизации архивных технологий:

- комплектование архивных фондов, экспертиза ценности и отбора документальных материалов;
- хранение архивной информации в распределенной базе данных;
- учет и обеспечение сохранности документов архивного фонда;
- автоматизация научно-справочного аппарата;
- подготовка к изданию архивных справочников и документальных материалов;
- планирование, учет деятельности и отчетность архивных учреждений.

Концепция ААТ предполагает прежде всего комплексный подход к внедрению компьютерной техники и автоматизированных технологий и рассматривается как совокупность методов и процессов обработки информации по комплектованию, экспертизе ценности, учету, сохранности, поиску документов, управлению архивным делом.

В традиционной технологии работа со справочниками и документами архива возможна только в местах их хранения. При автоматизированной технологии потребителю доступны все информационные массивы независимо от их местоположения.

## СПИСОК ЛИТЕРАТУРЫ

1. Автоматизированные рабочие места управленческого персонала. Кишинев, 1990.
2. Бойко В.В., Савинков В.М. Проектирование баз данных информационных систем. М., 1989.
3. Гарскова И.М. Базы и банки данных в исторических исследованиях. М., 1994.
4. Информационные системы в управлении производством. М., 1973.
5. Косташ И.В. АСУ и административная информационная система. Кишинев, 1989.
6. Криницкий Б.В., Мионов Г.А., Фролов Г.Д. Автоматизированные информационные системы. М., 1982.
7. Куликовский Л.Ф., Мотов В.В. Теоретические основы информационных процессов. М., 1987.
8. Лутфуллаев Х.С. Информационное обеспечение принятия решений в органах государственного управления. Ташкент, 1991.
9. Моисеева Н.К., Анискин Ю.П. Современное предприятие: Конкурентоспособность, маркетинг, обновление. М., 1993. Т.1,2.
10. Никитин П.И. Информационно-поисковые системы: Теоретические основы информационного поиска. М., 1972.
11. Мескон и др. Основы менеджмента. М., 1991.
12. Основы управления. Учеб. пособие. М., 1986.
13. Организация управления промышленным производством. М., 1986.
14. Роберт М. Фалмер. Энциклопедия современного управления. М., 1992. Т.1,3,4.
15. Соколов А.В. Информационно-поисковые системы. М., 1981.
16. Смирнов Э.Д. Автоматизация управленческого труда на базе персональных компьютеров. М., 1989.
17. Фурсин Г.И. Теория и практика создания банков данных. Киев, 1987.
18. Храмов Ю.А. ГИДРОНИМИКОН – экспертная система по гидронимии Восточно-Европейской равнины // Информ. бюллетень Комиссии по применению математических методов и ЭВМ в исторических исследованиях. 1992. №5.
19. Четвериков В.Н., Ревунков Г.И., Самохвалов Э.Н. Базы и банки данных. М., 1987.

## СПИСОК СОКРАЩЕНИЙ

АИПС – автоматизированная информационно-поисковая система  
АИС – автоматизированная информационная система  
АИЦ ИНИОН – Автоматизированный информационный центр  
Института научной информации по общественным наукам  
АСНТИ ГАРФ – автоматизированная система научно-технической информации Государственного архивного фонда Российской Федерации  
ББК – библиотечно-библиографическая классификация  
БД – база данных  
БнД – банк данных  
ВИНИТИ – Всероссийский институт научной и технической информации  
ВНИИДАД – Всероссийский научно-исследовательский институт документоведения и архивного дела  
ВТ – вычислительная техника  
ГАРФ – Государственный архивный фонд Российской Федерации  
ГАСНТИ – государственная автоматизированная система научно-технической информации  
ДИПС – документальная информационно-поисковая система  
ИВС – информационно-вычислительные сети  
ИИС – интегральная информационная система  
ИНИОН РАН – Институт научной информации по общественным наукам Российской Академии наук  
ИРИ – избирательное распространение информации  
ИПС – информационно-поисковые системы  
ИПЯ – информационно-поисковый язык  
ИС – информационная система  
НИД – научно-информационная деятельность  
НСА – научно-справочный аппарат  
НТИ – научно-техническая информация  
ПК – персональный компьютер  
САЦНИОН – сеть автоматизированных центров научной информации по общественным наукам  
СИ – сигнальная информация  
СУБД – система управления базами данных  
УДК – универсальная десятичная классификация  
ФИПС – фактографическая информационно-поисковая система  
ЦНИОН АН Грузии – Центр научной информации по общественным наукам Академии наук Грузии  
ЦНИОН АН Латвии – Центр научной информации по общественным наукам Академии наук Латвии  
ЭС – экспертная система

**Людмила Николаевна Мазур**

**ИНФОРМАЦИОННЫЕ СИСТЕМЫ**  
**Теоретические проблемы**

**Учебное пособие**

**Редактор В.И.Попова**  
**Технический редактор Э.А.Максимова**  
**ЛР №020257 от 22.11.96**

---

Подписано в печать 18.03.97.      Формат 60х84 1/16.  
Бумага для множительных аппаратов. Печать офсетная. **176**  
Уч.-изд.л. 4,16.      Усл.печ.л. 4,0.      Тираж 500 экз.      Заказ

Уральский государственный университет им. А.М.Горького.  
Екатеринбург, пр. Ленина, 51

---

Типолаборатория, УрГУ. Екатеринбург, пр.Ленина, 51

